PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 38/00, 39/385, 45/00, 45/05, C07K 7/00, 14/00, 14/82 — A1

(11) International Publication Number: **WO 96/10413**

(43) International Publication Date: 11 April 1996 (11.04.96)

(21) International Application Number: PCT/US95/12463

(22) International Filing Date: 27 September 1995 (27.09.95)

(30) Priority Data:
08/315,961 30 September 1994 (30.09.94) US

(71) Applicant: LUDWIG INSTITUTE FOR CANCER RESEARCH [CH/US]; 1345 Avenue of the Americas, New York, NY 10105 (US).

(72) Inventors: VAN DER BRUGGEN, Pierre; UCL 7459, Avenue Hippocrate 74, B-1200 Brussels (BE). BOON-FALLEUR, Thierry; UCL 7459, Avenue Hippocrate 74, B-1200 Brussels (BE). VAN DEN EYNDE, Benoit; UCL 7459, Avenue Hippocrate 74, B-1200 Brussels (BE).

(74) Agent: HANSON, Norman, D.; Felfe & Lynch, 805 Third Avenue, New York, NY 10022 (US).

(81) Designated States: AU, CA, CN, FI, JP, NO, NZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: COMPOSITIONS CONTAINING TUMOR REJECTION ANTIGEN PRECURSORS OR TUMOR REJECTION ANTIGENS, AND AN ADJUVANT AND/OR GROWTH FACTOR

(57) Abstract

Compositions with improved immunogenic effect are disclosed. The compositions contain one or more peptides which, when complexed with a cell surface molecule, such as an MHC, HLA or B cell receptor, provoke an immune response. The compositions contain adjuvants such as saponins, which potentiate the immune response. Especially preferred are compositions which stimulate cytolytic T cell responses, such as peptides which satisfy criteria for MHC binding, such as peptides derived from tumor rejection antigen precursors, including MAGE, BAGE, and GAGE derived peptides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

COMPOSITIONS CONTAINING TUMOR REJECTION ANTIGEN PRECURSORS OR TUMOR REJECTION ANTIGENS, AND AN ADJUVANT AND/OR GROWTH FACTOR

RELATED APPLICATIONS

This application is a continuation in part of all of the following applications, each of which discuss the provocation of cytolytic T cells in the presence of complexing tumor rejection antigens and MHC/HLA molecules: PCT application PCT/US92/04354 (designating the United States); Serial No. 938,334, filed August 31, 1992; Serial No. 008,446, filed January 22, 1993; Serial No. 54,714, filed April 28, 1993; Serial No. 203,054, filed February 28, 1994; Serial No. 233,305, filed April 26, 1994; Serial No. 195,186, filed February 14, 1994; Serial No. 196,630, filed February 15, 1994; Serial No. 32,978, filed March 18, 1993; Serial Nos. 217,186; 217,187; and 217,188, all of which were filed on March 24, 1994; Serial No. 190,411, filed on April 1, 1994 and Serial No. 253,503, filed June 3, 1994.

FIELD OF THE INVENTION

This invention relates to compositions useful in the generation of an immune response against the classes of molecules referred to as tumor rejection antigen precursors ("TRAPs") and tumor rejection antigens ("TRAs"). The immune response includes, inter alia, the production of antibodies against the TRAPs and TRAs, as well as T cells specific for complexes of TRA and major histocompatibility molecules ("MHCs"). Such T cells and antibodies may be generated, e.g., in a mouse, rat, rabbit, sheep, goat or other non-human animal, and then used in diagnostic methods to identify tumor presence. The compositions may also be used, therapeutically, via administration to a subject afflicted with a cancerous condition or one where cell transformation has taken place, such as melanoma or dysplastic nevi, to provoke an

immune response against tumors, cancer cells, and transformed cells.

BACKGROUND AND PRIOR ART

The study of the recognition or lack of recognition of cancer cells by a host organism has proceeded in many different directions. Understanding of the field presumes some understanding of both basic immunology and oncology.

Early research on mouse tumors revealed that these displayed molecules which led to rejection of tumor cells when transplanted into syngeneic animals. These molecules are "recognized" by T-cells in the recipient animal, and provoke a cytolytic T-cell response with lysis of the transplanted cells. This evidence was first obtained with tumors induced in vitro by chemical carcinogens, such as methylcholanthrene. The antigens expressed by the tumors and which elicited the T-cell response were found to be different for each tumor. See Prehn, et al., J. Natl. Canc. Inst. 18: 769-778 (1957); Klein et al., Cancer Res. 20: 1561-1572 (1960); Gross, Cancer Res. 3: 326-333 (1943), Basombrio, Cancer Res. 30: 2458-2462 (1970) for general teachings on inducing tumors with chemical carcinogens and differences in cell surface antigens. This class of antigens has come to be known as "tumor specific transplantation antigens" or "TSTAs". Following the observation of the presentation of such antigens when induced by chemical carcinogens, similar results were obtained when tumors were induced in vitro via ultraviolet radiation. See Kripke, J. Natl. Canc. Inst. 53: 333-1336 (1974).

While T-cell mediated immune responses were observed for the types of tumor described supra, spontaneous tumors were thought to be generally non-immunogenic. These were therefore believed not to present antigens which provoked a response to the tumor in the tumor carrying subject. See Hewitt, et al., Brit. J. Cancer 33: 241-259 (1976).

The family of tum$^-$ antigen presenting cell lines are immunogenic variants obtained by mutagenesis of mouse tumor cells or cell lines, as described by Boon et al., J. Exp. Med.

152: 1184-1193 (1980), the disclosure of which is incorporated by reference. To elaborate, tum$^-$ antigens are obtained by mutating tumor cells which do not generate an immune response in syngeneic mice and will form tumors (i.e., "tum$^+$" cells). When these tum$^+$ cells are mutagenized, they are rejected by syngeneic mice, and fail to form tumors (thus "tum$^-$"). See Boon et al., Proc. Natl. Acad. Sci. USA 74: 272 (1977), the disclosure of which is incorporated by reference. Many tumor types have been shown to exhibit this phenomenon. See, e.g., Frost et al., Cancer Res. 43: 125 (1983).

It appears that tum$^-$ variants fail to form progressive tumors because they elicit an immune rejection process. The evidence in favor of this hypothesis includes the ability of "tum$^-$" variants of tumors, i.e., those which do not normally form tumors, to do so in mice with immune systems suppressed by sublethal irradiation, Van Pel et al., Proc. Natl, Acad. Sci. USA 76: 5282-5285 (1979); and the observation that intraperitoneally injected tum$^-$ cells of mastocytoma P815 multiply exponentially for 12-15 days, and then are eliminated in only a few days in the midst of an influx of lymphocytes and macrophages (Uyttenhove et al., J. Exp. Med. 152: 1175-1183 (1980)). Further evidence includes the observation that mice acquire an immune memory which permits them to resist subsequent challenge to the same tum$^-$ variant, even when immunosuppressive amounts of radiation are administered with the following challenge of cells (Boon et al., Proc. Natl, Acad. Sci. USA 74: 272-275 (1977); Van Pel et al., supra; Uyttenhove et al., supra). Later research found that when spontaneous tumors were subjected to mutagenesis, immunogenic variants were produced which did generate a response. Indeed, these variants were able to elicit an immune protective response against the original tumor. See Van Pel et al., J. Exp. Med. 157: 1992-2001 (1983). Thus, it has been shown that it is possible to elicit presentation of a so-called "tumor rejection antigen" in a tumor which is a target for a syngeneic rejection response. Similar results have been obtained when foreign genes have been transfected into

spontaneous tumors. See Fearson et al., Cancer Res. 48: 2975-1980 (1988) in this regard.

A class of antigens has been recognized which are presented on the surface of tumor cells and are recognized by cytotoxic T cells, leading to lysis. This class of antigens will be referred to as "tumor rejection antigens" or "TRAs" hereafter. TRAs may or may not elicit antibody responses. The extent to which these antigens have been studied, has been via cytolytic T cell characterization studies in vitro i.e., the study of the identification of the antigen by a particular cytolytic T cell ("CTL" hereafter) subset. The subset proliferates upon recognition of the presented tumor rejection antigen, and the cells presenting the antigen are lysed. Characterization studies have identified CTL clones which specifically lyse cells expressing the antigens. Examples of this work may be found in Levy et al., Adv. Cancer Res. 24: 1-59 (1977); Boon et al., J. Exp. Med. 152: 1184-1193 (1980); Brunner et al., J. Immunol. 124: 1627-1634 (1980); Maryanski et al., Eur. J. Immunol. 124: 1627-1634 (1980); Maryanski et al., Eur. J. Immunol. 12: 406-412 (1982); Palladino et al., Canc. Res. 47: 5074-5079 (1987). This type of analysis is required for other types of antigens recognized by CTLs, including major histocompatibility antigens, the male specific H-Y antigens, and a class of antigens, referred to as "tum-" antigens, and discussed herein.

A tumor exemplary of the subject matter described supra is known as P815. See DePlaen et al., Proc. Natl. Acad. Sci. USA 85: 2274-2278 (1988); Szikora et al., EMBO J 9: 1041-1050 (1990), and Sibille et al., J. Exp. Med. 172: 35-45 (1990), the disclosures of which are incorporated by reference. The P815 tumor is a mastocytoma, induced in a DBA/2 mouse with methylcholanthrene and cultured as both an in vitro tumor and a cell line. The P815 line has generated many tum$^-$ variants following mutagenesis, including variants referred to as P91A (DePlaen, supra), 35B (Szikora, supra), and P198 (Sibille, supra). In contrast to tumor rejection antigens - and this is a key distinction - the tum$^-$ antigens are only present after

the tumor cells are mutagenized. Tumor rejection antigens are present on cells of a given tumor without mutagenesis. Hence, with reference to the literature, a cell line can be tum^+, such as the line referred to as "P1", and can be provoked to produce tum^- variants. Since the tum^- phenotype differs from that of the parent cell line, one expects a difference in the DNA of tum^- cell lines as compared to their tum^+ parental lines, and this difference can be exploited to locate the gene of interest in tum^- cells. As a result, it was found that genes of tum^- variants such as P91A, 35B and P198 differ from their normal alleles by point mutations in the coding regions of the gene. See Szikora and Sibille, supra, and Lurquin et al., Cell 58: 293-303 (1989). This has proved not to be the case with the TRAs of this invention. These papers also demonstrated that peptides derived from the tum^- antigen are presented by the L^d molecule for recognition by CTLs. P91A is presented by L^d, P35 by D^d and P198 by K^d.

Prior patent applications PCT/US92/04354, U.S. Serial No. 807,043; 764,364; 728,838 and 705,702, all of which are incorporated by reference and U.S. Patent No. 5,342,774, describe inventions involving, inter alia, genes and other nucleic acid molecules which code for various TRAPs, which are in turn processed to tumor rejection antigen, or "TRAs". SEQ ID NOS: 1-26 which are a part of the subject application present sequences of genes coding for various TRAPs, and the TRA referred to hereafter as MZ2E, which is derived from MAGE-1 TRAP (SEQ ID NO: 26).

The genes are useful as a source for the isolated and purified tumor rejection antigen precursor and the TRA themselves, either of which can be used as an agent for treating the cancer for which the antigen is a "marker", as well as in various diagnostic and surveillance approaches to oncology, discussed infra. It is known, for example, that tum^- cells can be used to generate CTLs which lyse cells presenting different tum^- antigens as well as tum^+ cells. See, e.g., Maryanski et al., Eur. J. Immunol 12: 401 (1982); and Van den Eynde et al., Modern Trends in Leukemia IX (June

1990), the disclosures of which are incorporated by reference. The tumor rejection antigen precursor may be expressed in cells transfected by the gene, and then used to generate an immune response against a tumor of interest.

In the parallel case of human neoplasms, it has been observed that autologous mixed lymphocyte-tumor cell cultures ("MLTC" hereafter) frequently generate responder lymphocytes which lyse autologous tumor cells and do not lyse natural killer targets, autologous EBV-transformed B cells, or autologous fibroblasts (see Anichini et al., Immunol. Today 8: 385-389 (1987)). This response has been particularly well studied for melanomas, and MLTC have been carried out either with peripheral blood cells or with tumor infiltrating lymphocytes. Examples of the literature in this area including Knuth et al., Proc. Natl. Acad. Sci. USA 86: 2804-2802 (1984); Mukherji et al., J. Exp. Med. 158: 240 (1983); Hérin et all, Int. J. Canc. 39: 390-396 (1987); Topalian et al, J. Clin. Oncol 6: 839-853 (1988). Stable cytolytic T cell clones have been derived from MLTC responder cells, and these clones are specific for the tumor cells. See Mukherji et al., supra, Hérin et all, supra, Knuth et al., supra. The antigens recognized on tumor cells by these autologous CTLs do not appear to represent a cultural artifact, since they are found on tumor cells in vivo. Topalian et al., supra; Degiovanni et al., Eur. J. Immunol. 20: 1865-1868 (1990). These observations, coupled with the techniques used herein to isolate the genes for specific murine tumor rejection antigen precursors, have led to the isolation of nucleic acid sequences coding for tumor rejection antigen precursors of TRAs presented on human tumors. It is now possible to isolate the nucleic acid sequences which code for tumor rejection antigen precursors, including, but not being limited to those most characteristic of a particular tumor, with ramifications that are described infra.

Additional work has focused upon the presentation of TRAs by the class of molecules known as major histocompatibility complexes, or "MHCs". Human forms of these molecules are

"human leukocyte antigens" or "HLAs". This work has resulted in several unexpected discoveries regarding the field. Specifically in U.S. patent application Serial Number 938,334, now U.S. Patent No. _____ the disclosure of which is incorporated by reference, nonapeptides are taught which are presented by the HLA-A1 molecule. The reference teaches that given the known specificity of particular peptides for particular HLA molecules, one should expect particular peptides to bind one HLA molecule. These peptides, presented herein as SEQ ID NOS: 27-34 are also presented in Traversari et al., J. Exp. Med. 176: 1453-1457 (1992). This is important, because different individuals possess different HLA phenotypes. As a result, while identification of particular peptides or of particular motifs, and the peptides which are members thereof, as being partners for a specific HLA molecule has diagnostic and therapeutic ramifications, these are only relevant for individuals with that particular HLA phenotype. There is a need for further work in the area, because cellular abnormalities are not restricted to one particular HLA phenotype, and targeted therapy requires some knowledge of the phenotype of the abnormal cells at issue.

In U.S. Patent Application Serial Number 008,446, filed January 22, 1993 and incorporated by reference, the fact that the MAGE-1 expression product is processed to a second TRA is disclosed. This second TRA is presented by HLA-Cw* 1601 molecules. The disclosure shows that a given TRAP can yield a plurality of TRAs.

In U.S. Patent Application Serial Number 994,928, filed December 22, 1992, and incorporated by reference herein, tyrosinase is described as a tumor rejection antigen precursor. This is a well known molecule as per Kwon, U.S. Patent No. 4,898,814. This reference discloses that a molecule which is produced by some normal cells (e.g., melanocytes), is processed in tumor cells to yield a tumor rejection antigen that is presented by HLA-A2 molecules. The peptide presented thereby is described in U.S. Application Serial No. 54,714, filed April 28, 1993. SEQ ID NO: 35 sets

forth this peptide. Additional tyrosinase derived peptides presented by HLA molecules are set forth in Serial Nos. 203,054, and 233,305 filed February 28, 1994 and April 26, 1994 and incorporated by reference (SEQ ID NOS: 36-41).

Other peptides which are TRAs are described in additional patent applications. U.S. Patent Application Serial No. 195,186, filed February 14, 1994, and incorporated by reference herein, sets forth three peptides (SEQ ID NOS: 42-44 herein), which are derived from MAGE-1 and which complex with HLA-Cw* 1601. Serial No. 196,630, filed February 15, 1994, discloses an unrelated tumor rejection antigen precursor, the so-called "BAGE" gene, and peptides derived therefrom, which are processed and then presented by HLA-Cw* 1601. These are set forth as SEQ ID NOS: 45-48, and this application is incorporated by reference. SEQ ID NO: 48 is the tumor rejection antigen. Additional coding sequences for a tumor rejection antigen precursor are set forth in Serial No. 32,978, filed March 18, 1993 and incorporated by reference. These are included herein as SEQ ID NOS: 49 and 50. A more extended sequence for this gene is set forth in Serial No. 272,351, filed July 8, 1994 incorporated by reference, and is SEQ ID NO: 51. In Serial No. 96,039, filed July 22, 1993, the sequence of tumor rejection antigen precursor GAGE is set forth. See SEQ ID NO: 52 for this information.

A series of peptides which provoke lysis by cytolytic T cells when presented by MHC molecules are set forth in Serial No. 217,186, Serial No. 08/217,188, and Serial No. 217,187, all filed on March 24, 1994, and all of which are incorporated by reference herein. The first of these applications discloses MAGE-3 derived peptides presented by HLA-A2. Five peptides are of interest. These are repeated here as SEQ ID NOS: 53-57. The second application presents 11 sequences derived from MAGE-2, believed to complex with HLA-A2.1 molecules (SEQ ID NOS: 58-68). The last of these applications discloses two additional peptides (SEQ ID NOS: 69 and 70) derived from MAGE-3 which complex to HLA-A2. Serial No. 190,411, filed April 1, 1994 and incorporated by reference,

sets forth three peptides (SEQ ID NO: 71-73), derived from MAGE-1, which are immunogenic in that they provoke production of antibodies in a host animal to which they have been administered. Seial No. 253,503, filed June 3, 1994 and incorporated by reference, teaches a further tumor rejection antigen precursor gene (SEQ ID NO:74), and a peptide, derived therefrom (SEQ ID NO:75), whch is presented by HLA-B44 molecules. Further in the application of Coulie, Ikeda and Boon-Falleur, filed concurrently, a sequence coding for a tumor rejection antigen precursor known as DAGE (SEQ ID NO:76) is set forth. DAGE is found almost universally on tumor cells, and only on testis cells with respect to normal cell expression. This makes it especially useful for cancer diagnosis and in the applications disclosed herein. The above listing should not be presumed to be exhaustive of the TRAP and TRA literature, but is presented to show its diversity and the fact that these materials not only provoke T cell proliferation, but also stimulate production of antibodies. It is well known that antibody producing cells can be used as a source to produce hybridomas, which in turn produce monoclonal antibodies. Thus, when the term "antibodies" is used herein, it encompasses both polyclonal and monoclonal antibodies.

The parent applications to the present case, including Serial No. 142,368 and Serial No. 190,411, both discuss the usefulness of combining TRAPs or TRAs with various materials as adjuvants, to produce vaccines, immunogenic compositions, etc. Adjuvants, broadly defined, are substances which promote immune responses. Frequently, the adjuvant of choice if Freund's complete adjuvant, or killed B. pertussis organisms, used in combination with alum precipitated antigen. A general discussion of adjuvants is provided in Goding, Monoclonal Antibodies: Principles & Practice (Second edition, 1986), at pages 61-63, which are incorporated by reference herein. Goding notes, however, that when the antigen of interest is of low molecular weight, or is poorly immunogenic, coupling to an immunogenic carrier is recommended. Such molecules, according

to Goding, generally have molecular weights below about 1000. Among the carriers suggested by Goding, at page 283, are keyhole limpet hemocyanin, bovine serum albumin, ovalbumin, and fowl immunoglobulin.

What is problematic about such carriers, however, is that frequently they are also immunogenic themselves. Thus, the immune response may be a general one, with part, most, or all of it being directed against the carrier molecule rather than the immunogen itself.

Exemplary of developments in the art as they relate to adjuvants is U.S. Patent No. 5,057,540 to Kensil, et al, the disclosure of which is incorporated by reference herein. Kensil et al disclose the preparation of various saponin extracts, which are useful as adjuvants in immunogenic compositions. As natural products, the extracts are not completely defined. Kensil, et al do provide a complete and enabling disclosure for how various extracts, including QA-7, QA-19, and QA-21 (also referred to as QS-21) are prepared. Experiments are set forth in which bovine serum albumin ("BSA") was combined with various extracts (examples 8 and 9), and where feline leukemia virus recombinant glycoprotein "gp70RΔ was tested, following absorption to aluminum hydroxide (alum). The two immunogens tested, however, are expected to be immunogenic in their own right (gp70RΔ has a molecular weight of 70 kd, and serum albumin has about the same molecular weight). No experiments were carried out at all on molecules which should, per se, be considered to be poorly or even non-immunogenic, and thus would be expected to require the use of alum absorption or the use of haptenic carriers for provocation of a response.

In PCT Application WO9219758, which corresponds to defensive publication 7697275, which is incorporated by reference herein, an adjuvant referred to as "MTP-MF59" is disclosed. This adjuvant is used in connection with a Plasmodium falciparum protein, "Pfs-25-B". This combination is described as a transmission blocking vaccine. The P. falciparum protein is itself large enough to be immunogenic.

Thus, none of the art shows that the improved adjuvants can be used in combination with presumptively non-immunogenic proteins and peptides to yield immunologically effective compositions. This is especially true for TRAP and TRA molecules, as outlined supra.

It has now been found, surprisingly, that compositions comprising tumor rejection antigen precursors or tumor rejection antigens can be made which, when administered to a subject animal, provoke an immunogenic response. In especially preferred embodiments the immunogenic portion of the composition consists of TRAP or, more preferably TRA molecules, of one or more types, and an adjuvant. Especially preferred are compositions where the adjuvant is QS21, as is disclosed in the Kensil, et al patent, incorporated by reference supra.

The immunogens of this invention consist of TRAPs or TRAs, meaning that they do not include haptens, carriers, precipitated alum, or any of the materials normally associated with materials which are or are expected to be poorly immunogenic. In especially preferred embodiments, the compositions consist essentially of the immunogen and the adjuvant.

The invention is described in greater detail in the disclosure which follows.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Example

The following example demonstrates the use of peptide MZ2E (SEQ ID NO: 26) in connection with adjuvant QS-21. It is to be understood, however, that one may routinely substitute proteins or peptides (the first component of the composition) and adjuvants (the second component of the composition). The unexpected effect of the combination of MZ2E and QS21 should occur in the alternate combinations, i.e., when other peptides are used.

Subjects tested are patients with stage IV melanoma or high risk stage III malignant melanoma. Stage IV patients have a median survival time of one year following diagnosis,

and only a 15% chance of long-term survival (Balch, et al, Cutaneous Melanoma, J.b. Lippincott, Philadelphia, 1992). The standard therapy for these patients includes treatment with decarbazine or drug combinations with decarbazine; however, response rates only range from 8-25%, and there is no evidence of the treatment extending survival. Balch et al, supra. Patients with high risk Stage III melanoma (pT4 thick truncal primary tumors or extremity melanomas, with five or more positive regional lymph nodes) have a median survival of 1-2 years following onset, and a 19% chance of long term survival. Balch, et al, supra.

Twelve patients are used in the study, all of whom have either Stage IV or high risk Stage III malignant melanoma in accordance with the criteria provided supra, confirmed by review of histology.

The patients fulfil the following additional criteria:

(i) complete recovery from surgery;

(ii) no chemotherapy or immunotherapy, within the four weeks preceding immunization;

(iii) expected survival of at least three months;

(iv) Karnofsky Performance Status of 60 or more;

(v) Laboratory results as follows:

granulocytes	≥	2,500/min^3
lymphocytes	≥	700/min^3
platelets	≥	100,000/min^3
serum creatinine	≤	2.0 mg/100 ml
serum bilirubin	≤	2.0 mg/100 ml

(vi) for MZ2E immunization, patients are positive for HLA-A1;

(vii) for MZ2E immunization, the patients' tumors express MAGE-1;

(viii) they be over 19 years old and be capable of giving informed consent in writing.

Any potential subjects who satisfy any of the following criteria are excluded:

(i) clinically significant heart disease;

(ii) other serious illness, such as active infection requiring antibiotics, or bleeding disorders;

(iii) treatment with antihistamines, N-SAIDS, or corticosteroids;

(iv) immunodeficiency, splenectomy, or radiation therapy to the spleen;

(v) pregnancy or lactation;

(vi) women of childbearing age who do not use effective methods of contraception.

All subjects are treated as outpatients. They are immunized, subcutaneously, with MZ2E (30 ug or 300 ug) and QS21 (100 ug), in 0.3 ml phosphate buffered saline, pH 7.4. Six patients receive 30 ug of peptide, and six receive 300 ug. The first injection is into the deltoid region of the anterior aspect of the thigh, and the site of injection is changed with follow up injections. No injections are given into limbs where draining lymph nodes have been surgically removed or irradiated.

Injections are given at day 1, and then at days 8, 15, 22, and 57. Patients are monitored over a course of 12 weeks, unless intervention is required (as determined by the investigator). Any patients who show stabilization or tumor response remain in the study until disease progression is evidenced. Patients may also be removed from the study or receive different doses of the peptide, if toxicity reactions are observed.

The patients show response as follows. In a complete response, all signs, symptoms, biochemical and imaging evidence of tumor disappear for a period of at least 30 days. In a partial response, there is a decrease in size of all measurable tumors of at least 50% of the sum of products of the greatest and perpendicular diameters for at least 30 days, without the appearance of new lesions or progression of any new lesions. In a minor response, there is a decrease in size of all measurable tumors of at least 25% of the sum of products of greatest and perpendicular diameters, for at least 30 days, without appearance of new lesions or progression of

any lesion. In stable disease, there is a change of less than 25% of the sum of products of greatest and perpendicular diameters, for at least 30 days, without the appearance of new lesions or progression of lesions. In progression of the disease, there is appearance of new tumors, or an increase in the size of measurable tumors of at least 50% of the sum of the product of greatest and perpendicular diameter.

The foregoing example demonstrates a composition comprising an amount of a tumor rejection antigen, i.e., MZ2E, and an adjuvant, i.e., QS21, and the use of the composition in the in vivo treatment of cancer (i.e., melanoma). The tumor rejection antigen is used in an amount sufficient to provoke an immune response against tumor cells which present it on their surface.

The compositions of the invention comprise any tumor rejection antigen precursor ("TRAP") or tumor rejection antigen ("TRA"), in combination with a pharmaceutically acceptable adjuvant. Preferred embodiments of the invention utilize the TRAPs and TRAs discussed supra and set forth in SEQ ID NO: 65, as well as the adjuvants described in the Background section.

As will be seen from the foregoing discussion, an important aspect of the invention is stimulation of proliferation of T cells. This can be an initial stimulation or an augmentation of a prior stimulation. In particular, it is desirable to stimulate cytolytic T cells which present peptides, such as the TRAs described herein, on their surfaces. The cytolytic T cells recognize these complexes of MHC and peptide, bind thereto via their receptor, and proliferate. They also lyse the recognized cells. This response can be used not only in vivo, but in vitro, as it is well established that cytolytic T cells specific for particular complexes of MHC and peptide are present in the blood of subjects who have experienced cell transformation. By contacting a blood sample of individuals in vitro with the peptide of interest and cells which present the MHC molecule of interest, any cytolytic T cells in the blood sample will

expand, i.e., proliferate. This proliferation an be measured via any of the well known assays designed therefor. Especially preferred are the radioactive chromium (^{51}Cr) release assay, and the measurement of release of tumor necrosis factor (TNF).

The compositions are also useful as stimulators of B cell proliferation, or antibody production. Again, it is well known that B cells produce antibodies, and the size of their targets are well within the sizes of the tumor rejection antigens, and certainly the tumor rejection antigen precursors. As with T cells, the stimulation may be "ab initio", or an augmentation of a prior response, in vitro or in vivo.

The amount of TRAP or TRA used will vary, depending upon the purpose of the immunization and the subject to which it is administered. For example, in the case of generating murine antibodies which can then be used, e.g., to diagnose for the presence of cancer cells presenting a TRA, the amount of protein or peptide may be less than that used in a course of in vivo therapy, such as that described in the example, supra. In general, a preferred dose can range from about 1 ug to about 750 ug of protein or peptide per dose. In a preferred embodiment, the range is from about 10 ug to about 500 ug. Most preferably, anywhere from about 30 ug to about 300 ug per dose may be used. Of course, in the context of the therapeutic aspect of the invention, the investigator will modify the dose, as a six month old infant will require dosing different from a full grown man, e.g. The mode of administration may vary, with preferred forms being oral, subcutaneous, intramuscular, intravenous and intraperitoneal adminstration.

The choice of TRAP or TRA protein or peptide in the composition will depend upon parameters determinable by the artisan. It is art recognized, for example, that different TRAs are presented by the various MHC molecules. As such, if a subject is typed, using well known techniques, as presenting HLA-A2 molecules on the surface of tumor cells, one will use

a TRA presented by HLA-A2 molecules rather than one presented by, e.g., HLA-Cw* 1601. Similarly, using techniques such as polymerase chain reaction ("PCR"), lysis studies, and other assay methodologies which are well known in the art, one can determine which tumor rejection antigen precursor gene or genes are being expressed by a subject patient. This will lead to the decision as to what protein or peptide to use. Again, by way of example, if a subject's tumor cells are expressing MAGE-3 but not MAGE-1, the peptide used in immunization should be derived from MAGE-3, and not MAGE-1.

While the molecules discussed herein are referred to as "tumor" rejection antigens and "tumor" rejection antigen precursors, it is intended that their use, in a therapeutic and also a diagnostic context, extends beyond cancer per se. The art is familiar with pathological conditions, such as displastic nevis, which are not cancer per se, but where the cells of the afflicted individuals are in fact characterized by transformation. Any and all such conditions are within the intended ambit of the invention.

Other aspects of the invention will be clear to the skilled artisan and need not be reiterated here.

The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, it being recognized that various modifications are possible within the scope of the invention.

(1) GENERAL INFORMATION:

(i) APPLICANTS: Boon-Falleur, Thierry; van den Eynde, Benôit; van der Bruggen, Pierre

(ii) TITLE OF INVENTION: COMPOSITIONSCONTAININGTUMORREJECTION ANTIGEN PRECURSORS OR TUMOR REJECTION ANTIGENS, AND AN ADJUVANT AND/OR GROWTH FACTOR

(iii) NUMBER OF SEQUENCES: 76

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Felfe & Lynch

(B) STREET: 805 Third Avenue

(C) CITY: New York City

(D) STATE: New York

(E) COUNTRY: USA

(F) ZIP: 10022

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

(B) COMPUTER: IBM PS/2

(C) OPERATING SYSTEM: PC-DOS

(D) SOFTWARE: Wordperfect

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/315,961

(B) FILING DATE: 30-SEPTEMBER-1994

(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US92/04354

(B) FILING DATE: 22-MAY-1992

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/938,334

(B) FILING DATE: 31-AUGUST-1992

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/008,446

(B) FILING DATE: 22-JANUARY-1993

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/054,714

(B) FILING DATE: 28-APRIL-1993

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/203,054
(B) FILING DATE: 28-FEBRUARY-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/233,305
(B) FILING DATE: 26-APRIL-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/195,186
(B) FILING DATE: 14-FEBRUARY-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/196,630
(B) FILING DATE: 15-FEBRUARY-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/032,978
(B) FILING DATE: 18-MARCH-1993
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/217,186
(B) FILING DATE: 24-MARCH-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/217,187
(B) FILING DATE: 24-MARCH-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/217,188
(B) FILING DATE: 24-MARCH-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/190,411
(B) FILING DATE: 1-APRIL-1994
(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/253,503
(B) FILING DATE: 3-JUNE-1994
(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Hanson, Norman D.
(B) REGISTRATION NUMBER: 30,946
(C) REFERENCE/DOCKET NUMBER: LUD 5370
(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 688-9200
(B) TELEFAX: (212) 838-3884

(2) INFORMATION FOR SEQUENCE ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 462 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ACCACAGGAG AATGAAAAGA ACCCGGGACT CCCAAAGACG CTAGATGTGT GAAGATCCTG  60
ATCACTCATT GGGTGTCTGA GTTCTGCGAT ATTCATCCCT CAGCCAATGA GCTTACTGTT 120
CTCGTGGGGG GTTTGTGAGC CTTGGGTAGG AAGTTTTGCA AGTTCCGCCT ACAGCTCTAG 180
CTTGTGAATT TGTACCCTTT CACGTAAAAA AGTAGTCCAG AGTTTACTAC ACCCTCCCTC 240
CCCCCTCCCA CCTCGTGCTG TGCTGAGTTT AGAAGTCTTC CTTATAGAAG TCTTCCGTAT 300
AGAACTCTTC CGGAGGAAGG AGGGAGGACC CCCCCCCTTT GCTCTCCCAG CATGCATTGT 360
GTCAACGCCA TTGCACTGAG CTGGTCGAAG AAGTAAGCCG CTAGCTTGCG ACTCTACTCT 420
TATCTTAACT TAGCTCGGCT TCCTGCTGGT ACCCTTTGTG CC                     462
```

(2) INFORMATION FOR SEQUENCE ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 675 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
ATG TCT GAT AAC AAG AAA CCA GAC AAA GCC CAC AGT GGC TCA GGT GGT   48
Met Ser Asp Asn Lys Lys Pro Asp Lys Ala His Ser Gly Ser Gly Gly
                  5                  10                  15
GAC GGT GAT GGG AAT AGG TGC AAT TTA TTG CAC CGG TAC TCC CTG GAA   96
Asp Gly Asp Gly Asn Arg Cys Asn Leu Leu His Arg Tyr Ser Leu Glu
             20                  25                  30
GAA ATT CTG CCT TAT CTA GGG TGG CTG GTC TTC GCT GTT GTC ACA ACA  144
Glu Ile Leu Pro Tyr Leu Gly Trp Leu Val Phe Ala Val Val Thr Thr
```

```
                35                  40                  45
AGT TTT CTG GCG CTC CAG ATG TTC ATA GAC GCC CTT TAT GAG GAG CAG  192
Ser Phe Leu Ala Leu Gln Met Phe Ile Asp Ala Leu Tyr Glu Glu Gln
     50                  55                  60
TAT GAA AGG GAT GTG GCC TGG ATA GCC AGG CAA AGC AAG CGC ATG TCC  240
Tyr Glu Arg Asp Val Ala Trp Ile Ala Arg Gln Ser Lys Arg Met Ser
 65                  70                  75                  80
TCT GTC GAT GAG GAT GAA GAC GAT GAG GAT GAT GAG GAT GAC TAC TAC  288
Ser Val Asp Glu Asp Glu Asp Asp Glu Asp Asp Glu Asp Asp Tyr Tyr
                 85                  90                  95
GAC GAC GAG GAC GAC GAC GAC GAT GCC TTC TAT GAT GAT GAG GAT GAT  336
Asp Asp Glu Asp Asp Asp Asp Asp Ala Phe Tyr Asp Asp Glu Asp Asp
            100                 105                 110
GAG GAA GAA GAA TTG GAG AAC CTG ATG GAT GAT GAA TCA GAA GAT GAG  384
Glu Glu Glu Glu Leu Glu Asn Leu Met Asp Asp Glu Ser Glu Asp Glu
        115                 120                 125
GCC GAA GAA GAG ATG AGC GTG GAA ATG GGT GCC GGA GCT GAG GAA ATG  432
Ala Glu Glu Glu Met Ser Val Glu Met Gly Ala Gly Ala Glu Glu Met
    130                 135                 140
GGT GCT GGC GCT AAC TGT GCC TGT GTT CCT GGC CAT CAT TTA AGG AAG  480
Gly Ala Gly Ala Asn Cys Ala Cys Val Pro Gly His His Leu Arg Lys
145                 150                 155                 160
AAT GAA GTG AAG TGT AGG ATG ATT TAT TTC TTC CAC GAC CCT AAT TTC  528
Asn Glu Val Lys Cys Arg Met Ile Tyr Phe Phe His Asp Pro Asn Phe
                165                 170                 175
CTG GTG TCT ATA CCA GTG AAC CCT AAG GAA CAA ATG GAG TGT AGG TGT  576
Leu Val Ser Ile Pro Val Asn Pro Lys Glu Gln Met Glu Cys Arg Cys
            180                 185                 190
GAA AAT GCT GAT GAA GAG GTT GCA ATG GAA GAG GAA GAA GAA GAA GAG  624
Glu Asn Ala Asp Glu Glu Val Ala Met Glu Glu Glu Glu Glu Glu Glu
        195                 200                 210
GAG GAG GAG GAG GAA GAG GAA ATG GGA AAC CCG GAT GGC TTC TCA CCT  672
Glu Glu Glu Glu Glu Glu Glu Met Gly Asn Pro Asp Gly Phe Ser Pro
220                 225                 230                 235

TAG                                                             675
```

(2) INFORMATION FOR SEQUENCE ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 228 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GCATGCAGTT GCAAAGCCCA GAAGAAAGAA ATGGACAGCG GAAGAAGTGG TTGTTTTTTT 60

TTCCCCTTCA TTAATTTTCT AGTTTTTAGT AATCCAGAAA ATTTGATTTT GTTCTAAAGT 120

TCATTATGCA AAGATGTCAC CAACAGACTT CTGACTGCAT GGTGAACTTT CATATGATAC 180

ATAGGATTAC ACTTGTACCT GTTAAAAATA AAAGTTTGAC TTGCATAC 228

(2) INFORMATION FOR SEQUENCE ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1365 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ACCACAGGAG AATGAAAAGA ACCCGGGACT CCCAAAGACG CTAGATGTGT 50

GAAGATCCTG ATCACTCATT GGGTGTCTGA GTTCTGCGAT ATTCATCCCT 100

CAGCCAATGA GCTTACTGTT CTCGTGGGGG GTTTGTGAGC CTTGGGTAGG 150

AAGTTTTGCA AGTTCCGCCT ACAGCTCTAG CTTGTGAATT TGTACCCTTT 200

CACGTAAAAA AGTAGTCCAG AGTTTACTAC ACCCTCCCTC CCCCCTCCCA 250

CCTCGTGCTG TGCTGAGTTT AGAAGTCTTC CTTATAGAAG TCTTCCGTAT 300

AGAACTCTTC CGGAGGAAGG AGGGAGGACC CCCCCCCTTT GCTCTCCCAG 350

CATGCATTGT GTCAACGCCA TTGCACTGAG CTGGTCGAAG AAGTAAGCCG 400

CTAGCTTGCG ACTCTACTCT TATCTTAACT TAGCTCGGCT TCCTGCTGGT 450

ACCCTTTGTG CC 462

ATG TCT GAT AAC AAG AAA CCA GAC AAA GCC CAC AGT GGC TCA 504

GGT GGT GAC GGT GAT GGG AAT AGG TGC AAT TTA TTG CAC CGG 546
TAC TCC CTG GAA GAA ATT CTG CCT TAT CTA GGG TGG CTG GTC 588
TTC GCT GTT GTC ACA ACA AGT TTT CTG GCG CTC CAG ATG TTC 630
ATA GAC GCC CTT TAT GAG GAG CAG TAT GAA AGG GAT GTG GCC 672
TGG ATA GCC AGG CAA AGC AAG CGC ATG TCC TCT GTC GAT GAG 714
GAT GAA GAC GAT GAG GAT GAT GAG GAT GAC TAC TAC GAC GAC 756
GAG GAC GAC GAC GAC GAT GCC TTC TAT GAT GAT GAG GAT GAT 798
GAG GAA GAA GAA TTG GAG AAC CTG ATG GAT GAT GAA TCA GAA 840
GAT GAG GCC GAA GAA GAG ATG AGC GTG GAA ATG GGT GCC GGA 882
GCT GAG GAA ATG GGT GCT GGC GCT AAC TGT GCC TGT GTT CCT 924
GGC CAT CAT TTA AGG AAG AAT GAA GTG AAG TGT AGG ATG ATT 966
TAT TTC TTC CAC GAC CCT AAT TTC CTG GTG TCT ATA CCA GTG 1008
AAC CCT AAG GAA CAA ATG GAG TGT AGG TGT GAA AAT GCT GAT 1050
GAA GAG GTT GCA ATG GAA GAG GAA GAA GAA GAA GAG GAG GAG 1092
GAG GAG GAA GAG GAA ATG GGA AAC CCG GAT GGC TTC TCA CCT 1134
TAG 1137
GCATGCAGTT GCAAAGCCCA GAAGAAAGAA ATGGACAGCG GAAGAAGTGG 1187
TTGTTTTTTT TTCCCCTTCA TTAATTTTCT AGTTTTTAGT AATCCAGAAA 1237
ATTTGATTTT GTTCTAAAGT TCATTATGCA AAGATGTCAC CAACAGACTT 1287
CTGACTGCAT GGTGAACTTT CATATGATAC ATAGGATTAC ACTTGTACCT 1337
GTTAAAAATA AAAGTTTGAC TTGCATAC 1365

(2) INFORMATION FOR SEQUENCE ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4698 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ACCACAGGAG AATGAAAAGA ACCCGGGACT CCCAAAGACG CTAGATGTGT 50
GAAGATCCTG ATCACTCATT GGGTGTCTGA GTTCTGCGAT ATTCATCCCT 100
CAGCCAATGA GCTTACTGTT CTCGTGGGGG GTTTGTGAGC CTTGGGTAGG 150
AAGTTTTGCA AGTTCCGCCT ACAGCTCTAG CTTGTGAATT TGTACCCTTT 200

CACGTAAAAA AGTAGTCCAG AGTTTACTAC ACCCTCCCTC CCCCCTCCCA 250
CCTCGTGCTG TGCTGAGTTT AGAAGTCTTC CTTATAGAAG TCTTCCGTAT 300
AGAACTCTTC CGGAGGAAGG AGGGAGGACC CCCCCCCTTT GCTCTCCCAG 350
CATGCATTGT GTCAACGCCA TTGCACTGAG CTGGTCGAAG AAGTAAGCCG 400
CTAGCTTGCG ACTCTACTCT TATCTTAACT TAGCTCGGCT TCCTGCTGGT 450
ACCCTTTGTG CC 462
ATG TCT GAT AAC AAG AAA CCA GAC AAA GCC CAC AGT GGC TCA 504
GGT GGT GAC GGT GAT GGG AAT AGG TGC AAT TTA TTG CAC CGG 546
TAC TCC CTG GAA GAA ATT CTG CCT TAT CTA GGG TGG CTG GTC 588
TTC GCT GTT GTC ACA ACA AGT TTT CTG GCG CTC CAG ATG TTC 630
ATA GAC GCC CTT TAT GAG GAG CAG TAT GAA AGG GAT GTG GCC 672
TGG ATA GCC AGG CAA AGC AAG CGC ATG TCC TCT GTC GAT GAG 714
GAT GAA GAC GAT GAG GAT GAT GAG GAT GAC TAC TAC GAC GAC 756
GAG GAC GAC GAC GAC GAT GCC TTC TAT GAT GAT GAG GAT GAT 798
GAG GAA GAA GAA TTG GAG AAC CTG ATG GAT GAT GAA TCA GAA 840
GAT GAG GCC GAA GAA GAG ATG AGC GTG GAA ATG GGT GCC GGA 882
GCT GAG GAA ATG GGT GCT GGC GCT AAC TGT GCC T 916
GTGAGTAACC CGTGGTCTTT ACTCTAGATT CAGGTGGGGT GCATTCTTTA 966
CTCTTGCCCA CATCTGTAGT AAAGACCACA TTTTGGTTGG GGGTCATTGC 1016
TGGAGCCATT CCTGGCTCTC CTGTCCACGC CTATCCCCGC TCCTCCCATC 1066
CCCCACTCCT TGCTCCGCTC TCTTTCCTTT TCCCACCTTG CCTCTGGAGC 1116
TTCAGTCCAT CCTGCTCTGC TCCCTTTCCC CTTTGCTCTC CTTGCTCCCC 1166
TCCCCCTCGG CTCAACTTTT CGTGCCTTCT GCTCTCTGAT CCCCACCCTC 1216
TTCAGGCTTC CCCATTTGCT CCTCTCCCGA AACCCTCCCC TTCCTGTTCC 1266
CCTTTTCGCG CCTTTTCTTT CCTGCTCCCC TCCCCCTCCC TATTTACCTT 1316
TCACCAGCTT TGCTCTCCCT GCTCCCCTCC CCCTTTTGCA CCTTTTCTTT 1366
TCCTGCTCCC CTCCCCCTCC CCTCCCTGTT TACCCTTCAC CGCTTTTCCT 1416
CTACCTGCTT CCCTCCCCCT TGCTGCTCCC TCCCTATTTG CATTTTCGGG 1466
TGCTCCTCCC TCCCCCTCCC CCTCCCTCCC TATTTGCATT TTCGGGTGCT 1516
CCTCCCTCCC CCTCCCCAGG CCTTTTTTTT TTTTTTTTTT TTTTTTTTTT 1566
TTGGTTTTTC GAGACAGGGT TTCTCTTTGT ATCCCTGGCT GTCCTGGCAC 1616
TCACTCTGTA GACCAGGCTG GCCTCAAACT CAGAAATCTG CCTGCCTCTG 1666
CCTCCCAAAT GCTGGGATTA AAGGCTTGCA CCAGGACTGC CCCAGTGCAG 1716
GCCTTTCTTT TTTCTCCTCT CTGGTCTCCC TAATCCCTTT TCTGCATGTT 1766
AACTCCCCTT TTGGCACCTT TCCTTTACAG GACCCCCTCC CCCTCCCTGT 1816
TTCCCTTCCG GCACCCTTCC TAGCCCTGCT CTGTTCCCTC TCCCTGCTCC 1866
CCTCCCCCTC TTTGCTCGAC TTTTAGCAGC CTTACCTCTC CCTGCTTTCT 1916

GCCCCGTTCC CCTTTTTTGT GCCTTTCCTC CTGGCTCCCC TCCACCTTCC 1966
AGCTCACCTT TTTGTTTGTT TGGTTGTTTG GTTGTTTGGT TTGCTTTTTT 2016
TTTTTTTTTT GCACCTTGTT TTCCAAGATC CCCCTCCCCC TCCGGCTTCC 2066
CCTCTGTGTG CCTTTCCTGT TCCCTCCCCC TCGCTGGCTC CCCCTCCCTT 2116
TCTGCCTTTC CTGTCCCTGC TCCCTTCTCT GCTAACCTTT TAATGCCTTT 2166
CTTTTCTAGA CTCCCCCCTC CAGGCTTGCT GTTTGCTTCT GTGCACTTTT 2216
CCTGACCCTG CTCCCCTTCC CCTCCCAGCT CCCCCCTCTT TTCCCACCTC 2266
CCTTTCTCCA GCCTGTCACC CCTCCTTCTC TCCTCTCTGT TTCTCCCACT 2316
TCCTGCTTCC TTTACCCCTT CCCTCTCCCT ACTCTCCTCC CTGCCTGCTG 2366
GACTTCCTCT CCAGCCGCCC AGTTCCCTGC AGTCCTGGAG TCTTTCCTGC 2416
CTCTCTGTCC ATCACTTCCC CCTAGTTTCA CTTCCCTTTC ACTCTCCCCT 2466
ATGTGTCTCT CTTCCTATCT ATCCCTTCCT TTCTGTCCCC TCTCCTCTGT 2516
CCATCACCTC TCTCCTCCCT TCCCTTTCCT CTCTCTTCCA TTTTCTTCCA 2566
CCTGCTTCTT TACCCTGCCT CTCCCATTGC CCTCTTACCT TTATGCCCAT 2616
TCCATGTCCC CTCTCAATTC CCTGTCCCAT TGTGCTCCCT CACATCTTCC 2666
ATTTCCCTCT TTCTCCCTTA GCCTCTTCTT CCTCTTCTCT TGTATCTCCC 2716
TTCCCTTTGC TTCTCCCTCC TCCTTTCCCC TTCCCCTATG CCCTCTACTC 2766
TACTTGATCT TCTCTCCTCT CCACATACCC TTTTTCCTTT CCACCCTGCC 2816
CTTTGTCCCC AGACCCTACA GTATCCTGTG CACAGGAAGT GGGAGGTGCC 2866
ATCAACAACA AGGAGGCAAG AAACAGAGCA AAATCCCAAA ATCAGCAGGA 2916
AAGGCTGGAT GAAAATAAGG CCAGGTTCTG AGGACAGCTG GAATCTAGCC 2966
AAGTGGCTCC TATAACCCTA AGTACCAAGG GAGAAAGTGA TGGTGAAGTT 3016
CTTGATCCTT GCTGCTTCTT TTACATATGT TGGCACATCT TTCTCAAATG 3066
CAGGCCATGC TCCATGCTTG GCGCTTGCTC AGCGTGGTTA AGTAATGGGA 3116
GAATCTGAAA ACTAGGGGCC AGTGGTTTGT TTTGGGGACA AATTAGCACG 3166
TAGTGATATT TCCCCCTAAA AATTATAACA AACAGATTCA TGATTTGAGA 3216
TCCTTCTACA GGTGAGAAGT GGAAAAATTG TCACTATGAA GTTCTTTTTA 3266
GGCTAAAGAT ACTTGGAACC ATAGAAGCGT TGTTAAAATA CTGCTTTCTT 3316
TTGCTAAAAT ATTCTTTCTC ACATATTCAT ATTCTCCAG 3355
GT GTT CCT GGC CAT CAT TTA AGG AAG AAT GAA GTG AAG TGT 3396
AGG ATG ATT TAT TTC TTC CAC GAC CCT AAT TTC CTG GTG TCT 3438
ATA CCA GTG AAC CCT AAG GAA CAA ATG GAG TGT AGG TGT GAA 3480
AAT GCT GAT GAA GAG GTT GCA ATG GAA GAG GAA GAA GAA GAA 3522
GAG GAG GAG GAG GAG GAA GAG GAA ATG GGA AAC CCG GAT GGC 3564
TTC TCA CCT TAG 3576
GCATGCAGGT ACTGGCTTCA CTAACCAACC ATTCCTAACA TATGCCTGTA 3626
GCTAAGAGCA TCTTTTTAAA AAATATTATT GGTAAACTAA ACAATTGTTA 3676

TCTTTTTACA TTAATAAGTA TTAAATTAAT CCAGTATACA GTTTTAAGAA 3726
CCCTAAGTTA AACAGAAGTC AATGATGTCT AGATGCCTGT TCTTTAGATT 3776
GTAGTGAGAC TACTTACTAC AGATGAGAAG TTGTTAGACT CGGGAGTAGA 3826
GACCAGTAAA AGATCATGCA GTGAAATGTG GCCATGGAAA TCGCATATTG 3876
TTCTTATAGT ACCTTTGAGA CAGCTGATAA CAGCTGACAA AAATAAGTGT 3926
TTCAAGAAAG ATCACACGCC ATGGTTCACA TGCAAATTAT TATTTTGTCG 3976
TTCTGATTTT TTTCATTTCT AGACCTGTGG TTTTAAAGAG ATGAAAATCT 4026
CTTAAAATTT CCTTCATCTT TAATTTTCCT TAACTTTAGT TTTTTTCACT 4076
TAGAATTCAA TTCAAATTCT TAATTCAATC TTAATTTTTA GATTTCTTAA 4126
AATGTTTTTT AAAAAAAATG CAAATCTCAT TTTTAAGAGA TGAAAGCAGA 4176
GTAACTGGGG GGCTTAGGGA ATCTGTAGGG TTGCGGTATA GCAATAGGGA 4226
GTTCTGGTCT CTGAGAAGCA GTCAGAGAGA ATGGAAAACC AGGCCCTTGC 4276
CAGTAGGTTA GTGAGGTTGA TATGATCAGA TTATGGACAC TCTCCAAATC 4326
ATAAATACTC TAACAGCTAA GGATCTCTGA GGGAAACACA ACAGGGAAAT 4376
ATTTTAGTTT CTCCTTGAGA AACAATGACA AGACATAAAA TTGGCAAGAA 4426
AGTCAGGAGT GTATTCTAAT AAGTGTTGCT TATCTCTTAT TTTCTTCTAC 4476
AGTTGCAAAG CCCAGAAGAA AGAAATGGAC AGCGGAAGAA GTGGTTGTTT 4526
TTTTTTCCCC TTCATTAATT TTCTAGTTTT TAGTAATCCA GAAAATTTGA 4576
TTTTGTTCTA AAGTTCATTA TGCAAAGATG TCACCAACAG ACTTCTGACT 4626
GCATGGTGAA CTTTCATATG ATACATAGGA TTACACTTGT ACCTGTTAAA 4676
AATAAAAGTT TGACTTGCAT AC 4698

(2) INFORMATION FOR SEQUENCE ID NO: 6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Leu Pro Tyr Leu Gly Trp Leu Val Phe
5

(2) INFORMATION FOR SEQUENCE ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2419 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGATCCAGGC CCTGCCAGGA AAAATATAAG GGCCCTGCGT GAGAACAGAG 50
GGGGTCATCC ACTGCATGAG AGTGGGGATG TCACAGAGTC CAGCCCACCC 100
TCCTGGTAGC ACTGAGAAGC CAGGGCTGTG CTTGCGGTCT GCACCCTGAG 150
GGCCCGTGGA TTCCTCTTCC TGGAGCTCCA GGAACCAGGC AGTGAGGCCT 200
TGGTCTGAGA CAGTATCCTC AGGTCACAGA GCAGAGGATG CACAGGGTGT 250
GCCAGCAGTG AATGTTTGCC CTGAATGCAC ACCAAGGGCC CCACCTGCCA 300
CAGGACACAT AGGACTCCAC AGAGTCTGGC CTCACCTCCC TACTGTCAGT 350
CCTGTAGAAT CGACCTCTGC TGGCCGGCTG TACCCTGAGT ACCCTCTCAC 400
TTCCTCCTTC AGGTTTTCAG GGGACAGGCC AACCCAGAGG ACAGGATTCC 450
CTGGAGGCCA CAGAGGAGCA CCAAGGAGAA GATCTGTAAG TAGGCCTTTG 500
TTAGAGTCTC CAAGGTTCAG TTCTCAGCTG AGGCCTCTCA CACACTCCCT 550
CTCTCCCCAG GCCTGTGGGT CTTCATTGCC CAGCTCCTGC CCACACTCCT 600
GCCTGCTGCC CTGACGAGAG TCATCATGTC TCTTGAGCAG AGGAGTCTGC 650
ACTGCAAGCC TGAGGAAGCC CTTGAGGCCC AACAAGAGGC CCTGGGCCTG 700
GTGTGTGTGC AGGCTGCCAC CTCCTCCTCC TCTCCTCTGG TCCTGGGCAC 750
CCTGGAGGAG GTGCCCACTG CTGGGTCAAC AGATCCTCCC CAGAGTCCTC 800
AGGGAGCCTC CGCCTTTCCC ACTACCATCA ACTTCACTCG ACAGAGGCAA 850
CCCAGTGAGG GTTCCAGCAG CCGTGAAGAG GAGGGGCCAA GCACCTCTTG 900
TATCCTGGAG TCCTTGTTCC GAGCAGTAAT CACTAAGAAG GTGGCTGATT 950
TGGTTGGTTT TCTGCTCCTC AAATATCGAG CCAGGGAGCC AGTCACAAAG 1000
GCAGAAATGC TGGAGAGTGT CATCAAAAAT TACAAGCACT GTTTTCCTGA 1050
GATCTTCGGC AAAGCCTCTG AGTCCTTGCA GCTGGTCTTT GGCATTGACG 1100
TGAAGGAAGC AGACCCCACC GGCCACTCCT ATGTCCTTGT CACCTGCCTA 1150
GGTCTCTCCT ATGATGGCCT GCTGGGTGAT AATCAGATCA TGCCCAAGAC 1200
AGGCTTCCTG ATAATTGTCC TGGTCATGAT TGCAATGGAG GGCGGCCATG 1250
CTCCTGAGGA GGAAATCTGG GAGGAGCTGA GTGTGATGGA GGTGTATGAT 1300
GGGAGGGAGC ACAGTGCCTA TGGGGAGCCC AGGAAGCTGC TCACCCAAGA 1350

TTTGGTGCAG GAAAAGTACC TGGAGTACGG CAGGTGCCGG ACAGTGATCC 1400
CGCACGCTAT GAGTTCCTGT GGGGTCCAAG GGCCCTCGCT GAAACCAGCT 1450
ATGTGAAAGT CCTTGAGTAT GTGATCAAGG TCAGTGCAAG AGTTCGCTTT 1500
TTCTTCCCAT CCCTGCGTGA AGCAGCTTTG AGAGAGGAGG AAGAGGGAGT 1550
CTGAGCATGA GTTGCAGCCA AGGCCAGTGG GAGGGGGACT GGGCCAGTGC 1600
ACCTTCCAGG GCCGCGTCCA GCAGCTTCCC CTGCCTCGTG TGACATGAGG 1650
CCCATTCTTC ACTCTGAAGA GAGCGGTCAG TGTTCTCAGT AGTAGGTTTC 1700
TGTTCTATTG GGTGACTTGG AGATTTATCT TTGTTCTCTT TTGGAATTGT 1750
TCAAATGTTT TTTTTTAAGG GATGGTTGAA TGAACTTCAG CATCCAAGTT 1800
TATGAATGAC AGCAGTCACA CAGTTCTGTG TATATAGTTT AAGGGTAAGA 1850
GTCTTGTGTT TTATTCAGAT TGGGAAATCC ATTCTATTTT GTGAATTGGG 1900
ATAATAACAG CAGTGGAATA AGTACTTAGA AATGTGAAAA ATGAGCAGTA 1950
AAATAGATGA GATAAAGAAC TAAAGAAATT AAGAGATAGT CAATTCTTGC 2000
CTTATACCTC AGTCTATTCT GTAAAATTTT TAAAGATATA TGCATACCTG 2050
GATTTCCTTG GCTTCTTTGA GAATGTAAGA GAAATTAAAT CTGAATAAAG 2100
AATTCTTCCT GTTCACTGGC TCTTTTCTTC TCCATGCACT GAGCATCTGC 2150
TTTTTGGAAG GCCCTGGGTT AGTAGTGGAG ATGCTAAGGT AAGCCAGACT 2200
CATACCCACC CATAGGGTCG TAGAGTCTAG GAGCTGCAGT CACGTAATCG 2250
AGGTGGCAAG ATGTCCTCTA AAGATGTAGG GAAAAGTGAG AGAGGGGTGA 2300
GGGTGTGGGG CTCCGGGTGA GAGTGGTGGA GTGTCAATGC CCTGAGCTGG 2350
GGCATTTTGG GCTTTGGGAA ACTGCAGTTC CTTCTGGGGG AGCTGATTGT 2400
AATGATCTTG GGTGGATCC 2419

(2) INFORMATION FOR SEQUENCE ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5674 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-1 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CCCGGGGCAC CACTGGCATC CCTCCCCCTA CCACCCCCAA TCCCTCCCTT 50
TACGCCACCC ATCCAAACAT CTTCACGCTC ACCCCCAGCC CAAGCCAGGC 100
AGAATCCGGT TCCACCCCTG CTCTCAACCC AGGGAAGCCC AGGTGCCCAG 150
ATGTGACGCC ACTGACTTGA GCATTAGTGG TTAGAGAGAA GCGAGGTTTT 200
CGGTCTGAGG GGCGGCTTGA GATCGGTGGA GGGAAGCGGG CCCAGCTCTG 250
TAAGGAGGCA AGGTGACATG CTGAGGGAGG ACTGAGGACC CACTTACCCC 300
AGATAGAGGA CCCCAAATAA TCCCTTCATG CCAGTCCTGG ACCATCTGGT 350
GGTGGACTTC TCAGGCTGGG CCACCCCCAG CCCCCTTGCT GCTTAAACCA 400
CTGGGGACTC GAAGTCAGAG CTCCGTGTGA TCAGGGAAGG GCTGCTTAGG 450
AGAGGGCAGC GTCCAGGCTC TGCCAGACAT CATGCTCAGG ATTCTCAAGG 500
AGGGCTGAGG GTCCCTAAGA CCCCACTCCC GTGACCCAAC CCCCACTCCA 550
ATGCTCACTC CCGTGACCCA ACCCCCTCTT CATTGTCATT CCAACCCCCA 600
CCCCACATCC CCCACCCCAT CCCTCAACCC TGATGCCCAT CCGCCCAGCC 650
ATTCCACCCT CACCCCCACC CCCACCCCCA CGCCCACTCC CACCCCCACC 700
CAGGCAGGAT CCGGTTCCCG CCAGGAAACA TCCGGGTGCC CGGATGTGAC 750
GCCACTGACT TGCGCATTGT GGGGCAGAGA GAAGCGAGGT TTCCATTCTG 800
AGGGACGGCG TAGAGTTCGG CCGAAGGAAC CTGACCCAGG CTCTGTGAGG 850
AGGCAAGGTG AGAGGCTGAG GGAGGACTGA GGACCCCGCC ACTCCAAATA 900
GAGAGCCCCA AATATTCCAG CCCCGCCCTT GCTGCCAGCC CTGGCCCACC 950
CGCGGGAAGA CGTCTCAGCC TGGGCTGCCC CCAGACCCCT GCTCCAAAAG 1000
CCTTGAGAGA CACCAGGTTC TTCTCCCCAA GCTCTGGAAT CAGAGGTTGC 1050
TGTGACCAGG GCAGGACTGG TTAGGAGAGG GCAGGGCACA GGCTCTGCCA 1100
GGCATCAAGA TCAGCACCCA AGAGGGAGGG CTGTGGGCCC CCAAGACTGC 1150
ACTCCAATCC CCACTCCCAC CCCATTCGCA TTCCCATTCC CCACCCAACC 1200
CCCATCTCCT CAGCTACACC TCCACCCCCA TCCCTACTCC TACTCCGTCA 1250
CCTGACCACC ACCCTCCAGC CCCAGCACCA GCCCCAACCC TTCTGCCACC 1300
TCACCCTCAC TGCCCCCAAC CCCACCCTCA TCTCTCTCAT GTGCCCCACT 1350
CCCATCGCCT CCCCCATTCT GGCAGAATCC GGTTTGCCCC TGCTCTCAAC 1400
CCAGGGAAGC CCTGGTAGGC CCGATGTGAA ACCACTGACT TGAACCTCAC 1450
AGATCTGAGA GAAGCCAGGT TCATTTAATG GTTCTGAGGG GCGGCTTGAG 1500
ATCCACTGAG GGGAGTGGTT TTAGGCTCTG TGAGGAGGCA AGGTGAGATG 1550
CTGAGGGAGG ACTGAGGAGG CACACACCCC AGGTAGATGG CCCCAAAATG 1600
ATCCAGTACC ACCCCTGCTG CCAGCCCTGG ACCACCCGGC CAGGACAGAT 1650
GTCTCAGCTG GACCACCCCC CGTCCCGTCC CACTGCCACT TAACCCACAG 1700
GGCAATCTGT AGTCATAGCT TATGTGACCG GGGCAGGGTT GGTCAGGAGA 1750
GGCAGGGCCC AGGCATCAAG GTCCAGCATC CGCCCGGCAT TAGGGTCAGG 1800
ACCCTGGGAG GGAACTGAGG GTTCCCCACC CACACCTGTC TCCTCATCTC 1850

CACCGCCACC CCACTCACAT TCCCATACCT ACCCCCTACC CCCAACCTCA 1900
TCTTGTCAGA ATCCCTGCTG TCAACCCACG GAAGCCACGG GAATGGCGGC 1950
CAGGCACTCG GATCTTGACG TCCCCATCCA GGGTCTGATG GAGGGAAGGG 2000
GCTTGAACAG GGCCTCAGGG GAGCAGAGGG AGGGCCCTAC TGCGAGATGA 2050
GGGAGGCCTC AGAGGACCCA GCACCCTAGG ACACCGCACC CCTGTCTGAG 2100
ACTGAGGCTG CCACTTCTGG CCTCAAGAAT CAGAACGATG GGGACTCAGA 2150
TTGCATGGGG GTGGGACCCA GGCCTGCAAG GCTTACGCGG AGGAAGAGGA 2200
GGGAGGACTC AGGGGACCTT GGAATCCAGA TCAGTGTGGA CCTCGGCCCT 2250
GAGAGGTCCA GGGCACGGTG GCCACATATG GCCCATATTT CCTGCATCTT 2300
TGAGGTGACA GGACAGAGCT GTGGTCTGAG AAGTGGGGCC TCAGGTCAAC 2350
AGAGGGAGGA GTTCCAGGAT CCATATGGCC CAAGATGTGC CCCCTTCATG 2400
AGGACTGGGG ATATCCCCGG CTCAGAAAGA AGGGACTCCA CACAGTCTGG 2450
CTGTCCCCTT TTAGTAGCTC TAGGGGGACC AGATCAGGGA TGGCGGTATG 2500
TTCCATTCTC ACTTGTACCA CAGGCAGGAA GTTGGGGGGC CCTCAGGGAG 2550
ATGGGGTCTT GGGGTAAAGG GGGGATGTCT ACTCATGTCA GGGAATTGGG 2600
GGTTGAGGAA GCACAGGCGC TGGCAGGAAT AAAGATGAGT GAGACAGACA 2650
AGGCTATTGG AATCCACACC CCAGAACCAA AGGGGTCAGC CCTGGACACC 2700
TCACCCAGGA TGTGGCTTCT TTTTCACTCC TGTTTCCAGA TCTGGGGCAG 2750
GTGAGGACCT CATTCTCAGA GGGTGACTCA GGTCAACGTA GGGACCCCCA 2800
TCTGGTCTAA AGACAGAGCG GTCCCAGGAT CTGCCATGCG TTCGGGTGAG 2850
GAACATGAGG GAGGACTGAG GGTACCCCAG GACCAGAACA CTGAGGGAGA 2900
CTGCACAGAA ATCAGCCCTG CCCCTGCTGT CACCCCAGAG AGCATGGGCT 2950
GGGCCGTCTG CCGAGGTCCT TCCGTTATCC TGGGATCATT GATGTCAGGG 3000
ACGGGGAGGC CTTGGTCTGA GAAGGCTGCG CTCAGGTCAG TAGAGGGAGC 3050
GTCCCAGGCC CTGCCAGGAG TCAAGGTGAG GACCAAGCGG GCACCTCACC 3100
CAGGACACAT TAATTCCAAT GAATTTTGAT ATCTCTTGCT GCCCTTCCCC 3150
AAGGACCTAG GCACGTGTGG CCAGATGTTT GTCCCCTCCT GTCCTTCCAT 3200
TCCTTATCAT GGATGTGAAC TCTTGATTTG GATTTCTCAG ACCAGCAAAA 3250
GGGCAGGATC CAGGCCCTGC CAGGAAAAAT ATAAGGGCCC TGCGTGAGAA 3300
CAGAGGGGGT CATCCACTGC ATGAGAGTGG GGATGTCACA GAGTCCAGCC 3350
CACCCTCCTG GTAGCACTGA GAAGCCAGGG CTGTGCTTGC GGTCTGCACC 3400
CTGAGGGCCC GTGGATTCCT CTTCCTGGAG CTCCAGGAAC CAGGCAGTGA 3450
GGCCTTGGTC TGAGACAGTA TCCTCAGGTC ACAGAGCAGA GGATGCACAG 3500
GGTGTGCCAG CAGTGAATGT TTGCCCTGAA TGCACACCAA GGGCCCCACC 3550
TGCCACAGGA CACATAGGAC TCCACAGAGT CTGGCCTCAC CTCCCTACTG 3600
TCAGTCCTGT AGAATCGACC TCTGCTGGCC GGCTGTACCC TGAGTACCCT 3650
CTCACTTCCT CCTTCAGGTT TTCAGGGGAC AGGCCAACCC AGAGGACAGG 3700

ATTCCCTGGA GGCCACAGAG GAGCACCAAG GAGAAGATCT GTAAGTAGGC 3750
CTTTGTTAGA GTCTCCAAGG TTCAGTTCTC AGCTGAGGCC TCTCACACAC 3800
TCCCTCTCTC CCCAGGCCTG TGGGTCTTCA TTGCCCAGCT CCTGCCCACA 3850
CTCCTGCCTG CTGCCCTGAC GAGAGTCATC 3880
ATG TCT CTT GAG CAG AGG AGT CTG CAC TGC AAG CCT GAG GAA 3922
GCC CTT GAG GCC CAA CAA GAG GCC CTG GGC CTG GTG TGT GTG 3964
CAG GCT GCC ACC TCC TCC TCC TCT CCT CTG GTC CTG GGC ACC 4006
CTG GAG GAG GTG CCC ACT GCT GGG TCA ACA GAT CCT CCC CAG 4048
AGT CCT CAG GGA GCC TCC GCC TTT CCC ACT ACC ATC AAC TTC 4090
ACT CGA CAG AGG CAA CCC AGT GAG GGT TCC AGC AGC CGT GAA 4132
GAG GAG GGG CCA AGC ACC TCT TGT ATC CTG GAG TCC TTG TTC 4174
CGA GCA GTA ATC ACT AAG AAG GTG GCT GAT TTG GTT GGT TTT 4216
CTG CTC CTC AAA TAT CGA GCC AGG GAG CCA GTC ACA AAG GCA 4258
GAA ATG CTG GAG AGT GTC ATC AAA AAT TAC AAG CAC TGT TTT 4300
CCT GAG ATC TTC GGC AAA GCC TCT GAG TCC TTG CAG CTG GTC 4342
TTT GGC ATT GAC GTG AAG GAA GCA GAC CCC ACC GGC CAC TCC 4384
TAT GTC CTT GTC ACC TGC CTA GGT CTC TCC TAT GAT GGC CTG 4426
CTG GGT GAT AAT CAG ATC ATG CCC AAG ACA GGC TTC CTG ATA 4468
ATT GTC CTG GTC ATG ATT GCA ATG GAG GGC GGC CAT GCT CCT 4510
GAG GAG GAA ATC TGG GAG GAG CTG AGT GTG ATG GAG GTG TAT 4552
GAT GGG AGG GAG CAC AGT GCC TAT GGG GAG CCC AGG AAG CTG 4594
CTC ACC CAA GAT TTG GTG CAG GAA AAG TAC CTG GAG TAC GGC 4636
AGG TGC CGG ACA GTG ATC CCG CAC GCT ATG AGT TCC TGT GGG 4678
GTC CAA GGG CCC TCG CTG AAA CCA GCT ATG TGA 4711
AAGTCCTTGA GTATGTGATC AAGGTCAGTG CAAGAGTTC 4750
GCTTTTTCTT CCCATCCCTG CGTGAAGCAG CTTTGAGAGA GGAGGAAGAG 4800
GGAGTCTGAG CATGAGTTGC AGCCAAGGCC AGTGGGAGGG GGACTGGGCC 4850
AGTGCACCTT CCAGGGCCGC GTCCAGCAGC TTCCCCTGCC TCGTGTGACA 4900
TGAGGCCCAT TCTTCACTCT GAAGAGAGCG GTCAGTGTTC TCAGTAGTAG 4950
GTTTCTGTTC TATTGGGTGA CTTGGAGATT TATCTTTGTT CTCTTTTGGA 5000
ATTGTTCAAA TGTTTTTTTT TAAGGGATGG TTGAATGAAC TTCAGCATCC 5050
AAGTTTATGA ATGACAGCAG TCACACAGTT CTGTGTATAT AGTTTAAGGG 5100
TAAGAGTCTT GTGTTTTATT CAGATTGGGA AATCCATTCT ATTTTGTGAA 5150
TTGGGATAAT AACAGCAGTG GAATAAGTAC TTAGAAATGT GAAAAATGAG 5200
CAGTAAAATA GATGAGATAA AGAACTAAAG AAATTAAGAG ATAGTCAATT 5250
CTTGCCTTAT ACCTCAGTCT ATTCTGTAAA ATTTTTAAAG ATATATGCAT 5300
ACCTGGATTT CCTTGGCTTC TTTGAGAATG TAAGAGAAAT TAAATCTGAA 5350

TAAAGAATTC TTCCTGTTCA CTGGCTCTTT TCTTCTCCAT GCACTGAGCA 5400
TCTGCTTTTT GGAAGGCCCT GGGTTAGTAG TGGAGATGCT AAGGTAAGCC 5450
AGACTCATAC CCACCCATAG GGTCGTAGAG TCTAGGAGCT GCAGTCACGT 5500
AATCGAGGTG GCAAGATGTC CTCTAAAGAT GTAGGGAAAA GTGAGAGAGG 5550
GGTGAGGGTG TGGGGCTCCG GGTGAGAGTG GTGGAGTGTC AATGCCCTGA 5600
GCTGGGGCAT TTTGGGCTTT GGGAAACTGC AGTTCCTTCT GGGGGAGCTG 5650
ATTGTAATGA TCTTGGGTGG ATCC 5674

(2) INFORMATION FOR SEQUENCE ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4157 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-2 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CCCATCCAGA TCCCCATCCG GGCAGAATCC GGTTCCACCC TTGCCGTGAA 50
CCCAGGGAAG TCACGGGCCC GGATGTGACG CCACTGACTT GCACATTGGA 100
GGTCAGAGGA CAGCGAGATT CTCGCCCTGA GCAACGGCCT GACGTCGGCG 150
GAGGGAAGCA GGCGCAGGCT CCGTGAGGAG GCAAGGTAAG ACGCCGAGGG 200
AGGACTGAGG CGGGCCTCAC CCCAGACAGA GGGCCCCCAA TTAATCCAGC 250
GCTGCCTCTG CTGCCGGGCC TGGACCACCC TGCAGGGGAA GACTTCTCAG 300
GCTCAGTCGC CACCACCTCA CCCCGCCACC CCCCGCCGCT TTAACCGCAG 350
GGAACTCTGG CGTAAGAGCT TTGTGTGACC AGGGCAGGGC TGGTTAGAAG 400
TGCTCAGGGC CCAGACTCAG CCAGGAATCA AGGTCAGGAC CCCAAGAGGG 450
GACTGAGGGC AACCCACCCC CTACCCTCAC TACCAATCCC ATCCCCCAAC 500
ACCAACCCCA CCCCCATCCC TCAAACACCA ACCCCACCCC CAAACCCCAT 550
TCCCATCTCC TCCCCCACCA CCATCCTGGC AGAATCCGGC TTTGCCCCTG 600
CAATCAACCC ACGGAAGCTC CGGGAATGGC GGCCAAGCAC GCGGATCCTG 650
ACGTTCACAT GTACGGCTAA GGGAGGGAAG GGGTTGGGTC TCGTGAGTAT 700
GGCCTTTGGG ATGCAGAGGA AGGGCCCAGG CCTCCTGGAA GACAGTGGAG 750
TCCTTAGGGG ACCCAGCATG CCAGGACAGG GGGCCCACTG TACCCCTGTC 800

TCAAACTGAG CCACCTTTTC ATTCAGCCGA GGGAATCCTA GGGATGCAGA 850
CCCACTTCAG GGGGTTGGGG CCCAGCCTGC GAGGAGTCAA GGGGAGGAAG 900
AAGAGGGAGG ACTGAGGGGA CCTTGGAGTC CAGATCAGTG GCAACCTTGG 950
GCTGGGGGAT CCTGGGCACA GTGGCCGAAT GTGCCCCGTG CTCATTGCAC 1000
CTTCAGGGTG ACAGAGAGTT GAGGGCTGTG GTCTGAGGGC TGGGACTTCA 1050
GGTCAGCAGA GGGAGGAATC CCAGGATCTG CCGGACCCAA GGTGTGCCCC 1100
CTTCATGAGG ACTCCCCATA CCCCCGGCCC AGAAAGAAGG GATGCCACAG 1150
AGTCTGGAAG TAAATTGTTC TTAGCTCTGG GGGAACCTGA TCAGGGATGG 1200
CCCTAAGTGA CAATCTCATT TGTACCACAG GCAGGAGGTT GGGGAACCCT 1250
CAGGGAGATA AGGTGTTGGT GTAAAGAGGA GCTGTCTGCT CATTTCAGGG 1300
GGTTCCCCCT TGAGAAAGGG CAGTCCCTGG CAGGAGTAAA GATGAGTAAC 1350
CCACAGGAGG CCATCATAAC GTTCACCCTA GAACCAAAGG GGTCAGCCCT 1400
GGACAACGCA CGTGGGGTAA CAGGATGTGG CCCCTCCTCA CTTGTCTTTC 1450
CAGATCTCAG GGAGTTGATG ACCTTGTTTT CAGAAGGTGA CTCAGTCAAC 1500
ACAGGGGCCC CTCTGGTCGA CAGATGCAGT GGTTCTAGGA TCTGCCAAGC 1550
ATCCAGGTGG AGAGCCTGAG GTAGGATTGA GGGTACCCCT GGGCCAGAAT 1600
GCAGCAAGGG GGCCCCATAG AAATCTGCCC TGCCCCTGCG GTTACTTCAG 1650
AGACCCTGGG CAGGGCTGTC AGCTGAAGTC CCTCCATTAT CTGGGATCTT 1700
TGATGTCAGG GAAGGGGAGG CCTTGGTCTG AAGGGGCTGG AGTCAGGTCA 1750
GTAGAGGGAG GGTCTCAGGC CCTGCCAGGA GTGGACGTGA GGACCAAGCG 1800
GACTCGTCAC CCAGGACACC TGGACTCCAA TGAATTTGAC ATCTCTCGTT 1850
GTCCTTCGCG GAGGACCTGG TCACGTATGG CCAGATGTGG GTCCCCTCTA 1900
TCTCCTTCTG TACCATATCA GGGATGTGAG TTCTTGACAT GAGAGATTCT 1950
CAAGCCAGCA AAAGGGTGGG ATTAGGCCCT ACAAGGAGAA AGGTGAGGGC 2000
CCTGAGTGAG CACAGAGGGG ACCCTCCACC CAAGTAGAGT GGGGACCTCA 2050
CGGAGTCTGG CCAACCCTGC TGAGACTTCT GGGAATCCGT GGCTGTGCTT 2100
GCAGTCTGCA CACTGAAGGC CCGTGCATTC CTCTCCCAGG AATCAGGAGC 2150
TCCAGGAACC AGGCAGTGAG GCCTTGGTCT GAGTCAGTGC CTCAGGTCAC 2200
AGAGCAGAGG GGACGCAGAC AGTGCCAACA CTGAAGGTTT GCCTGGAATG 2250
CACACCAAGG GCCCCACCCG CCCAGAACAA ATGGGACTCC AGAGGGCCTG 2300
GCCTCACCCT CCCTATTCTC AGTCCTGCAG CCTGAGCATG TGCTGGCCGG 2350
CTGTACCCTG AGGTGCCCTC CCACTTCCTC CTTCAGGTTC TGAGGGGGAC 2400
AGGCTGACAA GTAGGACCCG AGGCACTGGA GGAGCATTGA AGGAGAAGAT 2450
CTGTAAGTAA GCCTTTGTCA GAGCCTCCAA GGTTCAGTTC AGTTCTCACC 2500
TAAGGCCTCA CACACGCTCC TTCTCTCCCC AGGCCTGTGG GTCTTCATTG 2550
CCCAGCTCCT GCCCGCACTC CTGCCTGCTG CCCTGACCAG AGTCATC 2597
ATG CCT CTT GAG CAG AGG AGT CAG CAC TGC AAG CCT GAA GAA 2639

GGC CTT GAG GCC CGA GGA GAG GCC CTG GGC CTG GTG GGT GCG 2681
CAG GCT CCT GCT ACT GAG GAG CAG CAG ACC GCT TCT TCC TCT 2723
TCT ACT CTA GTG GAA GTT ACC CTG GGG GAG GTG CCT GCT GCC 2765
GAC TCA CCG AGT CCT CCC CAC AGT CCT CAG GGA GCC TCC AGC 2807
TTC TCG ACT ACC ATC AAC TAC ACT CTT TGG AGA CAA TCC GAT 2849
GAG GGC TCC AGC AAC CAA GAA GAG GAG GGG CCA AGA ATG TTT 2891
CCC GAC CTG GAG TCC GAG TTC CAA GCA GCA ATC AGT AGG AAG 2933
ATG GTT GAG TTG GTT CAT TTT CTG CTC CTC AAG TAT CGA GCC 2975
AGG GAG CCG GTC ACA AAG GCA GAA ATG CTG GAG AGT GTC CTC 3017
AGA AAT TGC CAG GAC TTC TTT CCC GTG ATC TTC AGC AAA GCC 3059
TCC GAG TAC TTG CAG CTG GTC TTT GGC ATC GAG GTG GTG GAA 3101
GTG GTC CCC ATC AGC CAC TTG TAC ATC CTT GTC ACC TGC CTG 3143
GGC CTC TCC TAC GAT GGC CTG CTG GGC GAC AAT CAG GTC ATG 3185
CCC AAG ACA GGC CTC CTG ATA ATC GTC CTG GCC ATA ATC GCA 3227
ATA GAG GGC GAC TGT GCC CCT GAG GAG AAA ATC TGG GAG GAG 3269
CTG AGT ATG TTG GAG GTG TTT GAG GGG AGG GAG GAC AGT GTC 3311
TTC GCA CAT CCC AGG AAG CTG CTC ATG CAA GAT CTG GTG CAG 3353
GAA AAC TAC CTG GAG TAC CGG CAG GTG CCC GGC AGT GAT CCT 3395
GCA TGC TAC GAG TTC CTG TGG GGT CCA AGG GCC CTC ATT GAA 3437
ACC AGC TAT GTG AAA GTC CTG CAC CAT ACA CTA AAG ATC GGT 3479
GGA GAA CCT CAC ATT TCC TAC CCA CCC CTG CAT GAA CGG GCT 3521
TTG AGA GAG GGA GAA GAG TGA 3542
GTCTCAGCAC ATGTTGCAGC CAGGGCCAGT GGGAGGGGGT CTGGGCCAGT 3592
GCACCTTCCA GGGCCCCATC CATTAGCTTC CACTGCCTCG TGTGATATGA 3642
GGCCCATTCC TGCCTCTTTG AAGAGAGCAG TCAGCATTCT TAGCAGTGAG 3692
TTTCTGTTCT GTTGGATGAC TTTGAGATTT ATCTTTCTTT CCTGTTGGAA 3742
TTGTTCAAAT GTTCCTTTTA ACAAATGGTT GGATGAACTT CAGCATCCAA 3792
GTTTATGAAT GACAGTAGTC ACACATAGTG CTGTTTATAT AGTTTAGGGG 3842
TAAGAGTCCT GTTTTTTATT CAGATTGGGA AATCCATTCC ATTTTGTGAG 3892
TTGTCACATA ATAACAGCAG TGGAATATGT ATTTGCCTAT ATTGTGAACG 3942
AATTAGCAGT AAAATACATG ATACAAGGAA CTCAAAAGAT AGTTAATTCT 3992
TGCCTTATAC CTCAGTCTAT TATGTAAAAT TAAAAATATG TGTATGTTTT 4042
TGCTTCTTTG AGAATGCAAA AGAAATTAAA TCTGAATAAA TTCTTCCTGT 4092
TCACTGGCTC ATTTCTTTAC CATTCACTCA GCATCTGCTC TGTGGAAGGC 4142
CCTGGTAGTA GTGGG 4157

(2) INFORMATION FOR SEQUENCE ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 662 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-21 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGATCCCCAT GGATCCAGGA AGAATCCAGT TCCACCCCTG CTGTGAACCC 50

AGGGAAGTCA CGGGGCCGGA TGTGACGCCA CTGACTTGCG CGTTGGAGGT 100

CAGAGAACAG CGAGATTCTC GCCCTGAGCA ACGGCCTGAC GTCGGCGGAG 150

GGAAGCAGGC GCAGGCTCCG TGAGGAGGCA AGGTAAGATG CCGAGGGAGG 200

ACTGAGGCGG GCCTCACCCC AGACAGAGGG CCCCCAATAA TCCAGCGCTG 250

CCTCTGCTGC CAGGCCTGGA CCACCCTGCA GGGGAAGACT TCTCAGGCTC 300

AGTCGCCACC ACCTCACCCC GCCACCCCCC GCCGCTTTAA CCGCAGGGAA 350

CTCTGGTGTA AGAGCTTTGT GTGACCAGGG CAGGGCTGGT TAGAAGTGCT 400

CAGGGCCCAG ACTCAGCCAG GAATCAAGGT CAGGACCCCA AGAGGGGACT 450

GAGGGTAACC CCCCCGCACC CCCACCACCA TTCCCATCCC CCAACACCAA 500

CCCCACCCCC ATCCCCCAAC ACCAAACCCA CCACCATCGC TCAAACATCA 550

ACGGCACCCC CAAACCCCGA TTCCCATCCC CACCCATCCT GGCAGAATCG 600

GAGCTTTGCC CCTGCAATCA ACCCACGGAA GCTCCGGGAA TGGCGGCCAA 650

GCACGCGGAT CC 662

(2) INFORMATION FOR SEQUENCE ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1640 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(ix) FEATURE:

(A) NAME/KEY: cDNA MAGE-3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GCCGCGAGGG AAGCCGGCCC AGGCTCGGTG AGGAGGCAAG GTTCTGAGGG 50
GACAGGCTGA CCTGGAGGAC CAGAGGCCCC CGGAGGAGCA CTGAAGGAGA 100
AGATCTGCCA GTGGGTCTCC ATTGCCCAGC TCCTGCCCAC ACTCCCGCCT 150
GTTGCCCTGA CCAGAGTCAT C 171
ATG CCT CTT GAG CAG AGG AGT CAG CAC TGC AAG CCT GAA GAA 213
GGC CTT GAG GCC CGA GGA GAG GCC CTG GGC CTG GTG GGT GCG 255
CAG GCT CCT GCT ACT GAG GAG CAG GAG GCT GCC TCC TCC TCT 297
TCT ACT CTA GTT GAA GTC ACC CTG GGG GAG GTG CCT GCT GCC 339
GAG TCA CCA GAT CCT CCC CAG AGT CCT CAG GGA GCC TCC AGC 381
CTC CCC ACT ACC ATG AAC TAC CCT CTC TGG AGC CAA TCC TAT 423
GAG GAC TCC AGC AAC CAA GAA GAG GAG GGG CCA AGC ACC TTC 465
CCT GAC CTG GAG TCC GAG TTC CAA GCA GCA CTC AGT AGG AAG 507
GTG GCC GAG TTG GTT CAT TTT CTG CTC CTC AAG TAT CGA GCC 549
AGG GAG CCG GTC ACA AAG GCA GAA ATG CTG GGG AGT GTC GTC 591
GGA AAT TGG CAG TAT TTC TTT CCT GTG ATC TTC AGC AAA GCT 633
TCC AGT TCC TTG CAG CTG GTC TTT GGC ATC GAG CTG ATG GAA 675
GTG GAC CCC ATC GGC CAC TTG TAC ATC TTT GCC ACC TGC CTG 717
GGC CTC TCC TAC GAT GGC CTG CTG GGT GAC AAT CAG ATC ATG 759
CCC AAG GCA GGC CTC CTG ATA ATC GTC CTG GCC ATA ATC GCA 801
AGA GAG GGC GAC TGT GCC CCT GAG GAG AAA ATC TGG GAG GAG 843
CTG AGT GTG TTA GAG GTG TTT GAG GGG AGG GAA GAC AGT ATG 885
TTG GGG GAT CCC AAG AAG CTG CTC ACC CAA CAT TTC GTG CAG 927
GAA AAC TAC CTG GAG TAC CGG CAG GTC CCC GGC AGT GAT CCT 969
GCA TGT TAT GAA TTC CTG TGG GGT CCA AGG GCC CTC GTT GAA 1011
ACC AGC TAT GTG AAA GTC CTG CAC CAT ATG GTA AAG ATC AGT 1053
GGA GGA CCT CAC ATT TCC TAC CCA CCC CTG CAT GAG TGG GTT 1095
TTG AGA GAG GGG GAA GAG TGA 1116
GTCTGAGCAC GAGTTGCAGC CAGGGCCAGT GGGAGGGGGT CTGGGCCAGT 1166
GCACCTTCCG GGGCCGCATC CCTTAGTTTC CACTGCCTCC TGTGACGTGA 1216
GGCCCATTCT TCACTCTTTG AAGCGAGCAG TCAGCATTCT TAGTAGTGGG 1266
TTTCTGTTCT GTTGGATGAC TTTGAGATTA TTCTTTGTTT CCTGTTGGAG 1316
TTGTTCAAAT GTTCCTTTTA ACGGATGGTT GAATGAGCGT CAGCATCCAG 1366
GTTTATGAAT GACAGTAGTC ACACATAGTG CTGTTTATAT AGTTTAGGAG 1416
TAAGAGTCTT GttTTTTACT CAAATTgGGA AATCCATTCC ATTTTGTGAA 1466

TTGTGACATA ATAATAGCAG TGGTAAAAGT ATTTGCTTAA AATTGTGAGC 1516
GAATTAGCAA TAACATACAT GAGATAACTC AAGAAATCAA AAGATAGTTG 1566
ATTCTTGCCT TGTACCTCAA TCTATTCTGT AAAATTAAAC AAATATGCAA 1616
ACCAGGATTT CCTTGACTTC TTTG 1640

(2) INFORMATION FOR SEQUENCE ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 943 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-31 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGATCCTCCA CCCCAGTAGA GTGGGGACCT CACAGAGTCT GGCCAACCCT 50
CCTGACAGTT CTGGGAATCC GTGGCTGCGT TTGCTGTCTG CACATTGGGG 100
GCCCGTGGAT TCCTCTCCCA GGAATCAGGA GCTCCAGGAA CAAGGCAGTG 150
AGGACTTGGT CTGAGGCAGT GTCCTCAGGT CACAGAGTAG AGGGGgCTCA 200
GATAGTGCCA ACGGTGAAGG TTTGCCTTGG ATTCAAACCA AGGGCCCCAC 250
CTGCCCCAGA ACACATGGAC TCCAGAGCGC CTGGCCTCAC CCTCAATACT 300
TTCAGTCCTG CAGCCTCAGC ATGCGCTGGC CGGATGTACC CTGAGGTGCC 350
CTCTCACTTC CTCCTTCAGG TTCTGAGGGG ACAGGCTGAC CTGGAGGACC 400
AGAGGCCCCC GGAGGAGCAC TGAAGGAGAA GATCTGTAAG TAAGCCTTTG 450
TTAGAGCCTC CAAGGTTCCA TTCAGTACTC AGCTGAGGTC TCTCACATGC 500
TCCCTCTCTC CCCAGGCCAG TGGGTCTCCA TTGCCCAGCT CCTGCCCACA 550
CTCCCGCCTG TTGCCCTGAC CAGAGTCATC 580
ATG CCT CTT GAG CAG AGG AGT CAG CAC TGC AAG CCT GAA GAA 622
GGC CTT GAG GCC CGA GGA GAg GCC CTG GGC CTG GTG GGT GCG 664
CAG GCT CCT GCT ACT GAG GAG CAG GAG GCT GCC TCC TCC TCT 706
TCT AGT GTA GTT GAA GTC ACC CTG GGG GAG GTG CCT GCT GCC 748
GAG TCA CCA GAT CCT CCC CAG AGT CCT CAG GGA GCC TCC AGC 790
CTC CCC ACT ACC ATG AAC TAC CCT CTC TGG AGC CAA TCC TAT 832
GAG GAC TCC AGC AAC CAA GAA GAG GAG GGG CCA AGC ACC TTC 874

CCT GAC CTG GAG TCT GAG TTC CAA GCA GCA CTC AGT AGG AAG 916
GTG GCC AAG TTG GTT CAT TTT CTG CTC 943

2) INFORMATION FOR SEQUENCE ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2531 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-4 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGATCCAGGC CCTGCCTGGA GAAATGTGAG GGCCCTGAGT GAACACAGTG 50
GGGATCATCC ACTCCATGAG AGTGGGGACC TCACAGAGTC CAGCCTACCC 100
TCTTGATGGC ACTGAGGGAC CGGGGCTGTG CTTACAGTCT GCACCCTAAG 150
GGCCCATGGA TTCCTCTCCT AGGAGCTCCA GGAACAAGGC AGTGAGGCCT 200
TGGTCTGAGA CAGTGTCCTC AGGTTACAGA GCAGAGGATG CACAGGCTGT 250
GCCAGCAGTG AATGTTTGCC CTGAATGCAC ACCAAGGGCC CCACCTGCCA 300
CAAGACACAT AGGACTCCAA AGAGTCTGGC CTCACCTCCC TACCATCAAT 350
CCTGCAGAAT CGACCTCTGC TGGCCGGCTA TACCCTGAGG TGCTCTCTCA 400
CTTCCTCCTT CAGGTTCTGA GCAGACAGGC CAACCGGAGA CAGGATTCCC 450
TGGAGGCCAC AGAGGAGCAC CAAGGAGAAG ATCTGTAAGT AAGCCTTTGT 500
TAGAGCCTCT AAGATTTGGT TCTCAGCTGA GGTCTCTCAC ATGCTCCCTC 550
TCTCCGTAGG CCTGTGGGTC CCCATTGCCC AGCTTTTGCC TGCACTCTTG 600
CCTGCTGCCC TGACCAGAGT CATC 624
ATG TCT TCT GAG CAG AAG AGT CAG CAC TGC AAG CCT GAG GAA 666
GGC GTT GAG GCC CAA GAA GAG GCC CTG GGC CTG GTG GGT GCA 708
CAG GCT CCT ACT ACT GAG GAG CAG GAG GCT GCT GTC TCC TCC 750
TCC TCT CCT CTG GTC CCT GGC ACC CTG GAG GAA GTG CCT GCT 792
GCT GAG TCA GCA GGT CCT CCC CAG AGT CCT CAG GGA GCC TCT 834
GCC TTA CCC ACT ACC ATC AGC TTC ACT TGC TGG AGG CAA CCC 876
AAT GAG GGT TCC AGC AGC CAA GAA GAG GAG GGG CCA AGC ACC 918

TCG CCT GAC GCA GAG TCC TTG TTC CGA GAA GCA CTC AGT AAC 960
AAG GTG GAT GAG TTG GCT CAT TTT CTG CTC CGC AAG TAT CGA 1002
GCC AAG GAG CTG GTC ACA AAG GCA GAA ATG CTG GAG AGA GTC 1044
ATC AAA AAT TAC AAG CGC TGC TTT CCT GTG ATC TTC GGC AAA 1086
GCC TCC GAG TCC CTG AAG ATG ATC TTT GGC ATT GAC GTG AAG 1128
GAA GTG GAC CCC GCC AGC AAC ACC TAC ACC CTT GTC ACC TGC 1170
CTG GGC CTT TCC TAT GAT GGC CTG CTG GGT AAT AAT CAG ATC 1212
TTT CCC AAG ACA GGC CTT CTG ATA ATC GTC CTG GGC ACA ATT 1254
GCA ATG GAG GGC GAC AGC GCC TCT GAG GAG GAA ATC TGG GAG 1296
GAG CTG GGT GTG ATG GGG GTG TAT GAT GGG AGG GAG CAC ACT 1338
GTC TAT GGG GAG CCC AGG AAA CTG CTC ACC CAA GAT TGG GTG 1380
CAG GAA AAC TAC CTG GAG TAC CGG CAG GTA CCC GGC AGT AAT 1422
CCT GCG CGC TAT GAG TTC CTG TGG GGT CCA AGG GCT CTG GCT 1464
GAA ACC AGC TAT GTG AAA GTC CTG GAG CAT GTG GTC AGG GTC 1506
AAT GCA AGA GTT CGC ATT GCC TAC CCA TCC CTG CGT GAA GCA 1548
GCT TTG TTA GAG GAG GAA GAG GGA GTC TGA 1578
GCATGAGTTG CAGCCAGGGC TGTGGGGAAG GGGCAGGGCT GGGCCAGTGC 1628
ATCTAACAGC CCTGTGCAGC AGCTTCCCTT GCCTCGTGTA ACATGAGGCC 1678
CATTCTTCAC TCTGTTTGAA GAAAATAGTC AGTGTTCTTA GTAGTGGGTT 1728
TCTATTTTGT TGGATGACTT GGAGATTTAT CTCTGTTTCC TTTTACAATT 1778
GTTGAAATGT TCCTTTTAAT GGATGGTTGA ATTAACTTCA GCATCCAAGT 1828
TTATGAATCG TAGTTAACGT ATATTGCTGT TAATATAGTT TAGGAGTAAG 1878
AGTCTTGTTT TTTATTCAGA TTGGGAAATC CGTTCTATTT TGTGAATTTG 1928
GGACATAATA ACAGCAGTGG AGTAAGTATT TAGAAGTGTG AATTCACCGT 1978
GAAATAGGTG AGATAAATTA AAAGATACTT AATTCCCGCC TTATGCCTCA 2028
GTCTATTCTG TAAAATTTAA AAATATATAT GCATACCTGG ATTTCCTTGG 2078
CTTCGTGAAT GTAAGAGAAA TTAAATCTGA ATAAATAATT CTTTCTGTTA 2128
ACTGGCTCAT TTCTTCTCTA TGCACTGAGC ATCTGCTCTG TGGAAGGCCC 2178
AGGATTAGTA GTGGAGATAC TAGGGTAAGC CAGACACACA CCTACCGATA 2228
GGGTATTAAG AGTCTAGGAG CGCGGTCATA TAATTAAGGT GACAAGATGT 2278
CCTCTAAGAT GTAGGGGAAA AGTAACGAGT GTGGGTATGG GGCTCCAGGT 2328
GAGAGTGGTC GGGTGTAAAT TCCCTGTGTG GGGCCTTTTG GGCTTTGGGA 2378
AACTGCATTT TCTTCTGAGG GATCTGATTC TAATGAAGCT TGGTGGGTCC 2428
AGGGCCAGAT TCTCAGAGGG AGAGGGAAAA GCCCAGATTG GAAAAGTTGC 2478
TCTGAGCAGT TCCTTTGTGA CAATGGATGA ACAGAGAGGA GCCTCTACCT 2528
GGG 2531

(2) INFORMATION FOR SEQUENCE ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2531 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-41 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GGATCCAGGC CCTGCCTGGA GAAATGTGAG GGCCCTGAGT GAACACAGTG 50

GGGATCATCC ACTCCATGAG AGTGGGGACC TCACAGAGTC CAGCCTACCC 100

TCTTGATGGC ACTGAGGGAC CGGGGCTGTG CTTACAGTCT GCACCCTAAG 150

GGCCCATGGA TTCCTCTCCT AGGAGCTCCA GGAACAAGGC AGTGAGGCCT 200

TGGTCTGAGA CAGTGTCCTC AGGTTACAGA GCAGAGGATG CACAGGCTGT 250

GCCAGCAGTG AATGTTTGCC CTGAATGCAC ACCAAGGGCC CCACCTGCCA 300

CAAGACACAT AGGACTCCAA AGAGTCTGGC CTCACCTCCC TACCATCAAT 350

CCTGCAGAAT CGACCTCTGC TGGCCGGCTA TACCCTGAGG TGCTCTCTCA 400

CTTCCTCCTT CAGGTTCTGA GCAGACAGGC CAACCGGAGA CAGGATTCCC 450

TGGAGGCCAC AGAGGAGCAC CAAGGAGAAG ATCTGTAAGT AAGCCTTTGT 500

TAGAGCCTCT AAGATTTGGT TCTCAGCTGA GGTCTCTCAC ATGCTCCCTC 550

TCTCCGTAGG CCTGTGGGTC CCCATTGCCC AGCTTTTGCC TGCACTCTTG 600

CCTGCTGCCC TGAGCAGAGT CATC 624

ATG TCT TCT GAG CAG AAG AGT CAG CAC TGC AAG CCT GAG GAA 666

GGC GTT GAG GCC CAA GAA GAG GCC CTG GGC CTG GTG GGT GCG 708

CAG GCT CCT ACT ACT GAG GAG CAG GAG GCT GCT GTC TCC TCC 750

TCC TCT CCT CTG GTC CCT GGC ACC CTG GAG GAA GTG CCT GCT 792

GCT GAG TCA GCA GGT CCT CCC CAG AGT CCT CAG GGA GCC TCT 834

GCC TTA CCC ACT ACC ATC AGC TTC ACT TGC TGG AGG CAA CCC 876

AAT GAG GGT TCC AGC AGC CAA GAA GAG GAG GGG CCA AGC ACC 918

TCG CCT GAC GCA GAG TCC TTG TTC CGA GAA GCA CTC AGT AAC 960

AAG GTG GAT GAG TTG GCT CAT TTT CTG CTC CGC AAG TAT CGA 1002

GCC AAG GAG CTG GTC ACA AAG GCA GAA ATG CTG GAG AGA GTC 1044

ATC AAA AAT TAC AAG CGC TGC TTT CCT GTG ATC TTC GGC AAA 1086
GCC TCC GAG TCC CTG AAG ATG ATC TTT GGC ATT GAC GTG AAG 1128
GAA GTG GAC CCC ACC AGC AAC ACC TAC ACC CTT GTC ACC TGC 1170
CTG GGC CTT TCC TAT GAT GGC CTG CTG GGT AAT AAT CAG ATC 1212
TTT CCC AAG ACA GGC CTT CTG ATA ATC GTC CTG GGC ACA ATT 1254
GCA ATG GAG GGC GAC AGC GCC TCT GAG GAG GAA ATC TGG GAG 1296
GAG CTG GGT GTG ATG GGG GTG TAT GAT GGG AGG GAG CAC ACT 1338
GTC TAT GGG GAG CCC AGG AAA CTG CTC ACC CAA GAT TGG GTG 1380
CAG GAA AAC TAC CTG GAG TAC CGG CAG GTA CCC GGC AGT AAT 1422
CCT GCG CGC TAT GAG TTC CTG TGG GGT CCA AGG GCT CTG GCT 1464
GAA ACC AGC TAT GTG AAA GTC CTG GAG CAT GTG GTC AGG GTC 1506
AAT GCA AGA GTT CGC ATT GCC TAC CCA TCC CTG CGT GAA GCA 1548
GCT TTG TTA GAG GAG GAA GAG GGA GTC TGA 1578
GCATGAGTTG CAGCCAGGGC TGTGGGGAAG GGGCAGGGCT GGGCCAGTGC 1628
ATCTAACAGC CCTGTGCAGC AGCTTCCCTT GCCTCGTGTA ACATGAGGCC 1678
CATTCTTCAC TCTGTTTGAA GAAAATAGTC AGTGTTCTTA GTAGTGGGTT 1728
TCTATTTTGT TGGATGACTT GGAGATTTAT CTCTGTTTCC TTTTACAATT 1778
GTTGAAATGT TCCTTTTAAT GGATGGTTGA ATTAACTTCA GCATCCAAGT 1828
TTATGAATCG TAGTTAACGT ATATTGCTGT TAATATAGTT TAGGAGTAAG 1878
AGTCTTGTTT TTTATTCAGA TTGGGAAATC CGTTCTATTT TGTGAATTTG 1928
GGACATAATA ACAGCAGTGG AGTAAGTATT TAGAAGTGTG AATTCACCGT 1978
GAAATAGGTG AGATAAATTA AAAGATACTT AATTCCCGCC TTATGCCTCA 2028
GTCTATTCTG TAAAATTTAA AAATATATAT GCATACCTGG ATTTCCTTGG 2078
CTTCGTGAAT GTAAGAGAAA TTAAATCTGA ATAAATAATT CTTTCTGTTA 2128
ACTGGCTCAT TTCTTCTCTA TGCACTGAGC ATCTGCTCTG TGGAAGGCCC 2178
AGGATTAGTA GTGGAGATAC TAGGGTAAGC CAGACACACA CCTACCGATA 2228
GGGTATTAAG AGTCTAGGAG CGCGGTCATA TAATTAAGGT GACAAGATGT 2278
CCTCTAAGAT GTAGGGGAAA AGTAACGAGT GTGGGTATGG GGCTCCAGGT 2328
GAGAGTGGTC GGGTGTAAAT TCCCTGTGTG GGGCCTTTTG GGCTTTGGGA 2378
AACTCCATTT TCTTCTGAGG GATCTGATTC TAATGAAGCT TGGTGGGTCC 2428
AGGGCCAGAT TCTCAGAGGG AGAGGGAAAA GCCCAGATTG GAAAAGTTGC 2478
TCTGAGCGGT TCCTTTGTGA CAATGGATGA ACAGAGAGGA GCCTCTACCT 2528
GGG 2531

(2) INFORMATION FOR SEQUENCE ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1068 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(ix) FEATURE:

(A) NAME/KEY: cDNA MAGE-4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

```
  G GGG CCA AGC ACC TCG CCT GAC GCA GAG TCC TTG TTC CGA      40
GAA GCA CTC AGT AAC AAG GTG GAT GAG TTG GCT CAT TTT CTG      82
CTC CGC AAG TAT CGA GCC AAG GAG CTG GTC ACA AAG GCA GAA     124
ATG CTG GAG AGA GTC ATC AAA AAT TAC AAG CGC TGC TTT CCT     166
GTG ATC TTC GGC AAA GCC TCC GAG TCC CTG AAG ATG ATC TTT     208
GGC ATT GAC GTG AAG GAA GTG GAC CCC GCC AGC AAC ACC TAC     250
ACC CTT GTC ACC TGC CTG GGC CTT TCC TAT GAT GGC CTG CTG     292
GGT AAT AAT CAG ATC TTT CCC AAG ACA GGC CTT CTG ATA ATC     334
GTC CTG GGC ACA ATT GCA ATG GAG GGC GAC AGC GCC TCT GAG     376
GAG GAA ATC TGG GAG GAG CTG GGT GTG ATG GGG GTG TAT GAT     418
GGG AGG GAG CAC ACT GTC TAT GGG GAG CCC AGG AAA CTG CTC     460
ACC CAA GAT TGG GTG CAG GAA AAC TAC CTG GAG TAC CGG CAG     502
GTA CCC GGC AGT AAT CCT GCG CGC TAT GAG TTC CTG TGG GGT     544
CCA AGG GCT CTG GCT GAA ACC AGC TAT GTG AAA GTC CTG GAG     586
CAT GTG GTC AGG GTC AAT GCA AGA GTT CGC ATT GCC TAC CCA     628
TCC CTG CGT GAA GCA GCT TTG TTA GAG GAG GAA GAG GGA GTC     670
TGAGCATGAG TTGCAGCCAG GGCTGTGGGG AAGGGGCAGG GCTGGGCCAG     720
TGCATCTAAC AGCCCTGTGC AGCAGCTTCC CTTGCCTCGT GTAACATGAG     770
GCCCATTCTT CACTCTGTTT GAAGAAAATA GTCAGTGTTC TTAGTAGTGG     820
GTTTCTATTT TGTTGGATGA CTTGGAGATT TATCTCTGTT TCCTTTTACA     870
ATTGTTGAAA TGTTCCTTTT AATGGATGGT TGAATTAACT TCAGCATCCA     920
AGTTTATGAA TCGTAGTTAA CGTATATTGC TGTTAATATA GTTTAGGAGT     970
AAGAGTCTTG TTTTTTATTC AGATTGGGAA ATCCGTTCTA TTTTGTGAAT    1020
TTGGGACATA ATAACAGCAG TGGAGTAAGT ATTTAGAAGT GTGAATTC      1068
```

(2) INFORMATION FOR SEQUENCE ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2226 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-5 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GGATCCAGGC CTTGCCAGGA GAAAGGTGAG GGCCCTGTGT GAGCACAGAG 50

GGGACCATTC ACCCCAAGAG GGTGGAGACC TCACAGATTC CAGCCTACCC 100

TCCTGTTAGC ACTGGGGGCC TGAGGCTGTG CTTGCAGTCT GCACCCTGAG 150

GGCCCATGCA TTCCTCTTCC AGGAGCTCCA GGAAACAGAC ACTGAGGCCT 200

TGGTCTGAGG CCGTGCCCTC AGGTCACAGA GCAGAGGAGA TGCAGACGTC 250

TAGTGCCAGC AGTGAACGTT TGCCTTGAAT GCACACTAAT GGCCCCCATC 300

GCCCCAGAAC ATATGGGACT CCAGAGCACC TGGCCTCACC CTCTCTACTG 350

TCAGTCCTGC AGAATCAGCC TCTGCTTGCT TGTGTACCCT GAGGTGCCCT 400

CTCACTTTTT CCTTCAGGTT CTCAGGGGAC AGGCTGACCA GGATCACCAG 450

GAAGCTCCAG AGGATCCCCA GGAGGCCCTA GAGGAGCACC AAAGGAGAAG 500

ATCTGTAAGT AAGCCTTTGT TAGAGCCTCC AAGGTTCAGT TTTTAGCTGA 550

GGCTTCTCAC ATGCTCCCTC TCTCTCCAGG CCAGTGGGTC TCCATTGCCC 600

AGCTCCTGCC CACACTCCTG CCTGTTGCGG TGACCAGAGT CGTC 644

ATG TCT CTT GAG CAG AAG AGT CAG CAC TGC AAG CCT GAG GAA 686

CTC CTC TGG TCC CAG GCA CCC TGG GGG AGG TGC CTG CTG CTG 728

GGT CAC CAG GTC CTC TCA AGA GTC CTC AGG GAG CCT CCG CCA 770

TCC CCA CTG CCA TCG ATT TCA CTC TAT GGA GGC AAT CCA TTA 812

AGG GCT CCA GCA ACC AAG AAG AGG AGG GGC CAA GCA CCT CCC 854

CTG ACC CAG AGT CTG TGT TCC GAG CAG CAC TCA GTA AGA AGG 896

TGG CTG ACT TGA 908

TTCATTTTCT GCTCCTCAAG TATTAAGTCA AGGAGCTGGT CACAAAGGCA 958

GAAATGCTGG AGAGCGTCAT CAAAAATTAC AAGCGCTGCT TTCCTGAGAT 1008

CTTCGGCAAA GCCTCCGAGT CCTTGCAGCT GGTCTTTGGC ATTGACGTGA 1058

AGGAAGCGGA CCCCACCAGC AACACCTACA CCCTTGTCAC CTGCCTGGGA 1108

CTCCTATGAT GGCCTGCTGG TTGATAATAA TCAGATCATG CCCAAGACGG 1158

GCCTCCTGAT AATCGTCTTG GGCATGATTG CAATGGAGGG CAAATGCGTC 1208

CCTGAGGAGA AAATCTGGGA GGAGCTGAGT GTGATGAAGG TGTATGTTGG 1258
GAGGGAGCAC AGTGTCTGTG GGGAGCCCAG GAAGCTGCTC ACCCAAGATT 1308
TGGTGCAGGA AAACTACCTG GAGTACCGGC AGGTGCCCAG CAGTGATCCC 1358
ATATGCTATG AGTTACTGTG GGGTCCAAGG GCACTCGCTG CTTGAAAGTA 1408
CTGGAGCACG TGGTCAGGGT CAATGCAAGA GTTCTCATTT CCTACCCATC 1458
CCTGCGTGAA GCAGCTTTGA GAGAGGAGGA AGAGGGAGTC TGAGCATGAG 1508
CTGCAGCCAG GGCCACTGCG AGGGGGGCTG GGCCAGTGCA CCTTCCAGGG 1558
CTCCGTCCAG TAGTTTCCCC TGCCTTAATG TGACATGAGG CCCATTCTTC 1608
TCTCTTTGAA GAGAGCAGTC AACATTCTTA GTAGTGGGTT TCTGTTCTAT 1658
TGGATGACTT TGAGATTTGT CTTTGTTTCC TTTTGGAATT GTTCAAATGT 1708
TTCTTTTAAT GGGTGGTTGA ATGAACTTCA GCATTCAAAT TTATGAATGA 1758
CAGTAGTCAC ACATAGTGCT GTTTATATAG TTTAGGAGTA AGAGTCTTGT 1808
TTTTTATTCA GATTGGGAAA TCCATTCCAT TTTGTGAATT GGGACATAGT 1858
TACAGCAGTG GAATAAGTAT TCATTTAGAA ATGTGAATGA GCAGTAAAAC 1908
TGATGACATA AAGAAATTAA AAGATATTTA ATTCTTGCTT ATACTCAGTC 1958
TATTCGGTAA AATTTTTTTT AAAAAATGTG CATACCTGGA TTTCCTTGGC 2008
TTCTTTGAGA ATGTAAGACA AATTAAATCT GAATAAATCA TTCTCCCTGT 2058
TCACTGGCTC ATTTATTCTC TATGCACTGA GCATTTGCTC TGTGGAAGGC 2108
CCTGGGTTAA TAGTGGAGAT GCTAAGGTAA GCCAGACTCA CCCCTACCCA 2158
CAGGGTAGTA AAGTCTAGGA GCAGCAGTCA TATAATTAAG GTGGAGAGAT 2208
GCCCTCTAAG ATGTAGAG 2226

(2) INFORMATION FOR SEQUENCE ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2305 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-51 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GGATCCAGGC CTTGCCAGGA GAAAGGTGAG GGCCCTGTGT GAGCACAGAG 50

GGGACCATTC ACCCCAAGAG GGTGGAGACC TCACAGATTC CAGCCTACCC 100
TCCTGTTAGC ACTGGGGGCC TGAGGCTGTG CTTGCAGTCT GCACCCTGAG 150
GGCCCATGCA TTCCTCTTCC AGGAGCTCCA GGAAACAGAC ACTGAGGCCT 200
TGGTCTGAGG CCGTGCCCTC AGGTCACAGA GCAGAGGAGA TGCAGACGTC 250
TAGTGCCAGC AGTGAACGTT TGCCTTGAAT GCACACTAAT GGCCCCCATC 300
GCCCCAGAAC ATATGGGACT CCAGAGCACC TGGCCTCACC CTCTCTACTG 350
TCAGTCCTGC AGAATCAGCC TCTGCTTGCT TGTGTACCCT GAGGTGCCCT 400
CTCACTTTTT CCTTCAGGTT CTCAGGGGAC AGGCTGACCA GGATCACCAG 450
GAAGCTCCAG AGGATCCCCA GGAGGCCCTA GAGGAGCACC AAAGGAGAAG 500
ATCTGTAAGT AAGCCTTTGT TAGAGCCTCC AAGGTTCAGT TTTTAGCTGA 550
GGCTTCTCAC ATGCTCCCTC TCTCTCCAGG CCAGTGGGTC TCCATTGCCC 600
AGCTCCTGCC CACACTCCTG CCTGTTGCGG TGACCAGAGT CGTC 644
ATG TCT CTT GAG CAG AAG AGT CAG CAC TGC AAG CCT GAG GAA 686
GGC CTT GAC ACC CAA GAA GAG CCC TGG GCC TGG TGG GTG TGC 728
AGG CTG CCA CTA CTG AGG AGC AGG AGG CTG TGT CCT CCT CCT 770
CTC CTC TGG TCC CAG GCA CCC TGG GGG AGG TGC CTG CTG CTG 812
GGT CAC CAG GTC CTC TCA AGA GTC CTC AGG GAG CCT CCG CCA 854
TCC CCA CTG CCA TCG ATT TCA CTC TAT GGA GGC AAT CCA TTA 896
AGG GCT CCA GCA ACC AAG AAG AGG AGG GGC CAA GCA CCT CCC 938
CTG ACC CAG AGT CTG TGT TCC GAG CAG CAC TCA GTA AGA AGG 980
TGG CTG ACT TGA 992
TTCATTTTCT GCTCCTCAAG TATTAAGTCA AGGAGCCGGT CACAAAGGCA 1042
GAAATGCTGG AGAGCGTCAT CAAAAATTAC AAGCGCTGCT TTCCTGAGAT 1092
CTTCGGCAAA GCCTCCGAGT CCTTGCAGCT GGTCTTTGGC ATTGACGTGA 1142
AGGAAGCGGA CCCCACCAGC AACACCTACA CCCTTGTCAC CTGCCTGGGA 1192
CTCCTATGAT GGCCTGGTGG TTTAATCAGA TCATGCCCAA GACGGGCCTC 1242
CTGATAATCG TCTTGGGCAT GATTGCAATG GAGGGCAAAT GCGTCCCTGA 1292
GGAGAAAATC TGGGAGGAGC TGGGTGTGAT GAAGGTGTAT GTTGGGAGGG 1342
AGCACAGTGT CTGTGGGGAG CCCAGGAAGC TGCTCACCCA AGATTTGGTG 1392
CAGGAAAACT ACCTGGAGTA CCGCAGGTGC CCAGCAGTGA TCCCATATGC 1442
TATGAGTTAC TGTGGGGTCC AAGGGCACTC GCTGCTTGAA AGTACTGGAG 1492
CACGTGGTCA GGGTCAATGC AAGAGTTCTC ATTTCCTACC CATCCCTGCA 1542
TGAAGCAGCT TTGAGAGAGG AGGAAGAGGG AGTCTGAGCA TGAGCTGCAG 1592
CCAGGGCCAC TGCGAGGGGG GCTGGGCCAG TGCACCTTCC AGGGCTCCGT 1642
CCAGTAGTTT CCCCTGCCTT AATGTGACAT GAGGCCCATT CTTCTCTCTT 1692
TGAAGAGAGC AGTCAACATT CTTAGTAGTG GGTTTCTGTT CTATTGGATG 1742
ACTTTGAGAT TTGTCTTTGT TTCCTTTTGG AATTGTTCAA ATGTTCCTTT 1792

TAATGGGTGG TTGAATGAAC TTCAGCATTC AAATTTATGA ATGACAGTAG 1842
TCACACATAG TGCTGTTTAT ATAGTTTAGG AGTAAGAGTC TTGTTTTTTA 1892
TTCAGATTGG GAAATCCATT CCATTTTGTG AATTGGGACA TAGTTACAGC 1942
AGTGGAATAA GTATTCATTT AGAAATGTGA ATGAGCAGTA AAACTGATGA 1992
GATAAAGAAA TTAAAAGATA TTTAATTCTT GCCTTATACT CAGTCTATTC 2042
GGTAAAATTT TTTTTTAAAA ATGTGCATAC CTGGATTTCC TTGGCTTCTT 2092
TGAGAATGTA AGACAAATTA AATCTGAATA AATCATTCTC CCTGTTCACT 2142
GGCTCATTTA TTCTCTATGC ACTGAGCATT TGCTCTGTGG AAGGCCCTGG 2192
GTTAATAGTG GAGATGCTAA GGTAAGCCAG ACTCACCCCT ACCCACAGGG 2242
TAGTAAAGTC TAGGAGCAGC AGTCATATAA TTAAGGTGGA GAGATGCCCT 2292
CTAAGATGTA GAG 2305

(2) INFORMATION FOR SEQUENCE ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 225 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-6 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TAT TTC TTT CCT GTG ATC TTC AGC AAA GCT TCC GAT TCC TTG 42
CAG CTG GTC TTT GGC ATC GAG CTG ATG GAA GTG GAC CCC ATC 84
GGC CAC GTG TAC ATC TTT GCC ACC TGC CTG GGC CTC TCC TAC 126
GAT GGC CTG CTG GGT GAC AAT CAG ATC ATG CCC AGG ACA GGC 168
TTC CTG ATA ATC ATC CTG GCC ATA ATC GCA AGA GAG GGC GAC 210
TGT GCC CCT GAG GAG 225

(2) INFORMATION FOR SEQUENCE ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1947 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: genomic DNA
(ix) FEATURE:
(A) NAME/KEY: MAGE-7 gene
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TGAATGGACA ACAAGGGCCC CACACTCCCC AGAACACAAG GGACTCCAGA 50
GAGCCCAGCC TCACCTTCCC TACTGTCAGT CCTGCAGCCT CAGCCTCTGC 100
TGGCCGGCTG TACCCTGAGG TGCCCTCTCA CTTCCTCCTT CAGGTTCTCA 150
GCGGACAGGC CGGCCAGGAG GTCAGAAGCC CCAGGAGGCC CCAGAGGAGC 200
ACCGAAGGAG AAGATCTGTA AGTAGGCCTT TGTTAGGGCC TCCAGGGCGT 250
GGTTCACAAA TGAGGCCCCT CACAAGCTCC TTCTCTCCCC AGATCTGTGG 300
GTTCCTCCCC ATCGCCCAGC TGCTGCCCGC ACTCCAGCCT GCTGCCCTGA 350
CCAGAGTCAT CATGTCTTCT GAGCAGAGGA GTCAGCACTG CAAGCCTGAG 400
GATGCCTTGA GGCCCAAGGA CAGGAGGCTC TGGGCCTGGT GGGTGCGCAG 450
GCTCCCGCCA CCGAGGAGCA CGAGGCTGCC TCCTCCTTCA CTCTGATTGA 500
AGGCACCCTG GAGGAGGTGC CTGCTGCTGG GTCCCCCAGT CCTCCCCTGA 550
GTCTCAGGGT TCCTCCTTTT CCCTGACCAT CAGCAACAAC ACTCTATGGA 600
GCCAATCCAG TGAGGGCACC AGCAGCCGGG AAGAGGAGGG GCCAACCACC 650
TAGACACACC CCGCTCACCT GGCGTCCTTG TTCCA 685
ATG GGA AGG TGG CTG AGT TGG TTC GCT TCC TGC TGC ACA AGT 727
ATC GAG TCA AGG AGC TGG TCA CAA AGG CAG AAA TGC TGG ACA 769
GTG TCA TCA AAA ATT ACA AGC ACT AGT TTC CTT GTG ATC TAT 811
GGC AAA GCC TCA GAG TGC ATG CAG GTG ATG TTT GGC ATT GAC 853
ATG AAG GAA GTG GAC CCC GCG GCC ACT CCT ACG TCC TTG TCA 895
CCT GCT TGG GCC TCT CCT ACA ATG GCC TGC TGG GTG ATG ATC 937
AGA GCA TGC CCG AGA CCG GCC TTC TGA 964
TTATGGTCTT GACCATGATC TTAATGGAGG GCCACTGTGC CCCTGAGGAG 1014
GCAATCTGGG AAGCGTTGAG TGTAATGGTG TATGATGGGA TGGAGCAGTT 1064
TCTTTGGGCA GCTGAGGAAG CTGCTCACCC AAGATTGGGT GCAGGAAAAC 1114
TACCTGCAAT ACCGCCAGGT GCCCAGCAGT GATCCCCCGT GCTACCAGTT 1164
CCTGTGGGGT CCAAGGGCCC TCATTGAAAC CAGCTATGTG AAAGTCCTGG 1214
AGTATGCAGC CAGGGTCAGT ACTAAAGAGA GCATTTCCTA CCCATCCCTG 1264

CATGAAGAGG CTTTGGGAGA GGAGGAAGAG GGAGTCTGAG CAGAAGTTGC 1314
AGCCAGGGCC AGTGGGGCAG ATTGGGGGAG GGCCTGGGCA GTGCACGTTC 1364
CACACATCCA CCACCTTCCC TGTCCTGTTA CATGAGGCCC ATTCTTCACT 1414
CTGTGTTTGA AGAGAGCAGT CAATGTTCTC AGTAGCGGGG AGTGTGTTGG 1464
GTGTGAGGGA ATACAAGGTG GACCATCTCT CAGTTCCTGT TCTCTTGGGC 1514
GATTTGGAGG TTTATCTTTG TTTCCTTTTG CAGTCGTTCA AATGTTCCTT 1564
TTAATGGATG GTGTAATGAA CTTCAACATT CATTTCATGT ATGACAGTAG 1614
GCAGACTTAC TGTTTTTTAT ATAGTTAAAA GTAAGTGCAT TGTTTTTTAT 1664
TTATGTAAGA AAATCTATGT TATTTCTTGA ATTGGGACAA CATAACATAG 1714
CAGAGGATTA AGTACCTTTT ATAATGTGAA AGAACAAAGC GGTAAAATGG 1764
GTGAGATAAA GAAATAAAGA AATTAAATTG GCTGGGCACG GTGGCTCACG 1814
CCTGTAATCC CAGCACTTTA GGAGGCAGAG GCACGGGGAT CACGAGGTCA 1864
GGAGATCGAG ACCATTCTGG CTAACACAGT GAAACACCAT CTCTATTAAA 1914
AATACAAAAC TTAGCCGGGC GTGGTGGCGG GTG 1947

(2) INFORMATION FOR SEQUENCE ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1810 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-8 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GAGCTCCAGG AACCAGGCTG TGAGGTCTTG GTCTGAGGCA GTATCTTCAA 50
TCACAGAGCA TAAGAGGCCC AGGCAGTAGT AGCAGTCAAG CTGAGGTGGT 100
GTTTCCCCTG TATGTATACC AGAGGCCCCT CTGGCATCAG AACAGCAGGA 150
ACCCCACAGT TCCTGGCCCT ACCAGCCCTT TTGTCAGTCC TGGAGCCTTG 200
GCCTTTGCCA GGAGGCTGCA CCCTGAGATG CCCTCTCAAT TTCTCCTTCA 250
GGTTCGCAGA GAACAGGCCA GCCAGGAGGT CAGGAGGCCC CAGAGAAGCA 300
CTGAAGAAGA CCTGTAAGTA GACCTTTGTT AGGGCATCCA GGGTGTAGTA 350
CCCAGCTGAG GCCTCTCACA CGCTTCCTCT CTCCCCAGGC CTGTGGGTCT 400
CAATTGCCCA GCTCCGGCCC ACACTCTCCT GCTGCCCTGA CCTGAGTCAT 450
C 451

ATG CTT CTT GGG CAG AAG AGT CAG CGC TAC AAG GCT GAG GAA 493
GGC CTT CAG GCC CAA GGA GAG GCA CCA GGG CTT ATG GAT GTG 535
CAG ATT CCC ACA GCT GAG GAG CAG AAG GCT GCA TCC TCC TCC 577
TCT ACT CTG ATC ATG GGA ACC CTT GAG GAG GTG ACT GAT TCT 619
GGG TCA CCA AGT CCT CCC CAG AGT CCT GAG GGT GCC TCC TCT 661
TCC CTG ACT GTC ACC GAC AGC ACT CTG TGG AGC CAA TCC GAT 703
GAG GGT TCC AGC AGC AAT GAA GAG GAG GGG CCA AGC ACC TCC 745
CCG GAC CCA GCT CAC CTG GAG TCC CTG TTC CGG GAA GCA CTT 787
GAT GAG AAA GTG GCT GAG TTA GTT CGT TTC CTG CTC CGC AAA 829
TAT CAA ATT AAG GAG CCG GTC ACA AAG GCA GAA ATG CTT GAG 871
AGT GTC ATC AAA AAT TAC AAG AAC CAC TTT CCT GAT ATC TTC 913
AGC AAA GCC TCT GAG TGC ATG CAG GTG ATC TTT GGC ATT GAT 955
GTG AAG GAA GTG GAC CCT GCC GGC CAC TCC TAC ATC CTT GTC 997
ACC TGC CTG GGC CTC TCC TAT GAT GGC CTG CTG GGT GAT GAT 1039
CAG AGT ACG CCC AAG ACC GGC CTC CTG ATA ATC GTC CTG GGC 1081
ATG ATC TTA ATG GAG GGC AGC CGC GCC CCG GAG GAG GCA ATC 1123
TGG GAA GCA TTG AGT GTG ATG GGG GCT GTA TGA 1156
TGGGAGGGAG CACAGTGTCT ATTGGAAGCT CAGGAAGCTG CTCACCCAAG 1206
AGTGGGTGCA GGAGAACTAC CTGGAGTACC GCCAGGCGCC CGGCAGTGAT 1256
CCTGTGCGCT ACGAGTTCCT GTGGGGTCCA AGGGCCCTTG CTGAAACCAG 1306
CTATGTGAAA GTCCTGGAGC ATGTGGTCAG GGTCAATGCA AGAGTTCGCA 1356
TTTCCTACCC ATCCCTGCAT GAAGAGGCTT TGGGAGAGGA GAAAGGAGTT 1406
TGAGCAGGAG TTGCAGCTAG GGCCAGTGGG GCAGGTTGTG GGAGGGCCTG 1456
GGCCAGTGCA CGTTCCAGGG CCACATCCAC CACTTTCCCT GCTCTGTTAC 1506
ATGAGGCCCA TTCTTCACTC TGTGTTTGAA GAGAGCAGTC ACAGTTCTCA 1556
GTAGTGGGGA GCATGTTGGG TGTGAGGGAA CACAGTGTGG ACCATCTCTC 1606
AGTTCCTGTT CTATTGGGCG ATTTGGAGGT TTATCTTTGT TTCCTTTTGG 1656
AATTGTTCCA ATGTTCCTTC TAATGGATGG TGTAATGAAC TTCAACATTC 1706
ATTTTATGTA TGACAGTAGA CAGACTTACT GCTTTTTATA TAGTTTAGGA 1756
GTAAGAGTCT TGCTTTTCAT TTATACTGGG AAACCCATGT TATTTCTTGA 1806
ATTC 1810

(2) INFORMATION FOR SEQUENCE ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1412 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: genomic DNA
(ix) FEATURE:
(A) NAME/KEY: MAGE-9 gene
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

TCTGAGACAG TGTCCTCAGG TCGCAGAGCA GAGGAGACCC AGGCAGTGTC 50
AGCAGTGAAG GTGAAGTGTT CACCCTGAAT GTGCACCAAG GGCCCCACCT 100
GCCCCAGCAC ACATGGGACC CCATAGCACC TGGCCCCATT CCCCCTACTG 150
TCACTCATAG AGCCTTGATC TCTGCAGGCT AGCTGCACGC TGAGTAGCCC 200
TCTCACTTCC TCCCTCAGGT TCTCGGGACA GGCTAACCAG GAGGACAGGA 250
GCCCCAAGAG GCCCCAGAGC AGCACTGACG AAGACCTGTA AGTCAGCCTT 300
TGTTAGAACC TCCAAGGTTC GGTTCTCAGC TGAAGTCTCT CACACACTCC 350
CTCTCTCCCC AGGCCTGTGG GTCTCCATCG CCCAGCTCCT GCCCACGCTC 400
CTGACTGCTG CCCTGACCAG AGTCATC 427
ATG TCT CTC GAG CAG AGG AGT CCG CAC TGC AAG CCT GAT GAA 469
GAC CTT GAA GCC CAA GGA GAG GAC TTG GGC CTG ATG GGT GCA 511
CAG GAA CCC ACA GGC GAG GAG GAG GAG ACT ACC TCC TCC TCT 553
GAC AGC AAG GAG GAG GAG GTG TCT GCT GCT GGG TCA TCA AGT 595
CCT CCC CAG AGT CCT CAG GGA GGC GCT TCC TCC TCC ATT TCC 637
GTC TAC TAC ACT TTA TGG AGC CAA TTC GAT GAG GGC TCC AGC 679
AGT CAA GAA GAG GAA GAG CCA AGC TCC TCG GTC GAC CCA GCT 721
CAG CTG GAG TTC ATG TTC CAA GAA GCA CTG AAA TTG AAG GTG 763
GCT GAG TTG GTT CAT TTC CTG CTC CAC AAA TAT CGA GTC AAG 805
GAG CCG GTC ACA AAG GCA GAA ATG CTG GAG AGC GTC ATC AAA 847
AAT TAC AAG CGC TAC TTT CCT GTG ATC TTC GGC AAA GCC TCC 889
GAG TTC ATG CAG GTG ATC TTT GGC ACT GAT GTG AAG GAG GTG 931
GAC CCC GCC GGC CAC TCC TAC ATC CTT GTC ACT GCT CTT GGC 973
CTC TCG TGC GAT AGC ATG CTG GGT GAT GGT CAT AGC ATG CCC 1015
AAG GCC GCC CTC CTG ATC ATT GTC CTG GGT GTG ATC CTA ACC 1057
AAA GAC AAC TGC GCC CCT GAA GAG GTT ATC TGG GAA GCG TTG 1099
AGT GTG ATG GGG GTG TAT GTT GGG AAG GAG CAC ATG TTC TAC 1141
GGG GAG CCC AGG AAG CTG CTC ACC CAA GAT TGG GTG CAG GAA 1183
AAC TAC CTG GAG TAC CGG CAG GTG CCC GGC AGT GAT CCT GCG 1225
CAC TAC GAG TTC CTG TGG GGT TCC AAG GCC CAC GCT GAA ACC 1267
AGC TAT GAG AAG GTC ATA AAT TAT TTG GTC ATG CTC AAT GCA 1309

AGA GAG CCC ATC TGC TAC CCA TCC CTT TAT GAA GAG GTT TTG 1351
GGA GAG GAG CAA GAG GGA GTC TGA 1375
GCACCAGCCG CAGCCGGGGC CAAAGTTTGT GGGGTCA 1412

(2) INFORMATION FOR SEQUENCE ID NO: 22:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 920 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: genomic DNA
(ix) FEATURE:
(A) NAME/KEY: MAGE-10 gene
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

ACCTGCTCCA GGACAAAGTG GACCCCACTG CATCAGCTCC ACCTACCCTA 50
CTGTCAGTCC TGGAGCCTTG GCCTCTGCCG GCTGCATCCT GAGGAGCCAT 100
CTCTCACTTC CTTCTTCAGG TTCTCAGGGG ACAGGGAGAG CAAGAGGTCA 150
AGAGCTGTGG GACACCACAG AGCAGCACTG AAGGAGAAGA CCTGTAAGTT 200
GGCCTTTGTT AGAACCTCCA GGGTGTGGTT CTCAGCTGTG GCCACTTACA 250
CCCTCCCTCT CTCCCCAGGC CTGTGGGTCC CCATCGCCCA AGTCCTGCCC 300
ACACTCCCAC CTGCTACCCT GATCAGAGTC ATC 333
ATG CCT CGA GCT CCA AAG CGT CAG CGC TGC ATG CCT GAA GAA 375
GAT CTT CAA TCC CAA AGT GAG ACA CAG GGC CTC GAG GGT GCA 417
CAG GCT CCC CTG GCT GTG GAG GAG GAT GCT TCA TCA TCC ACT 459
TCC ACC AGC TCC TCT TTT CCA TCC TCT TTT CCC TCC TCC TCC 501
TCT TCC TCC TCC TCC TCC TGC TAT CCT CTA ATA CCA AGC ACC 543
CCA GAG GAG GTT TCT GCT GAT GAT GAG ACA CCA AAT CCT CCC 585
CAG AGT GCT CAG ATA GCC TGC TCC TCC CCC TCG GTC GTT GCT 627
TCC CTT CCA TTA GAT CAA TCT GAT GAG GGC TCC AGC AGC CAA 669
AAG GAG GAG AGT CCA AGC ACC CTA CAG GTC CTG CCA GAC AGT 711
GAG TCT TTA CCC AGA AGT GAG ATA GAT GAA AAG GTG ACT GAT 753
TTG GTG CAG TTT CTG CTC TTC AAG TAT CAA ATG AAG GAG CCG 795
ATC ACA AAG GCA GAA ATA CTG GAG AGT GTC ATA AAA AAT TAT 837
GAA GAC CAC TTC CCT TTG TTG TTT AGT GAA GCC TCC GAG TGC 879

ATG CTG CTG GTC TTT GGC ATT GAT GTA AAG GAA GTG GAT CC 920

(2) INFORMATION FOR SEQUENCE ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1107 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: MAGE-11 gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

AGAGAACAGG CCAACCTGGA GGACAGGAGT CCCAGGAGAA CCCAGAGGAT 50
CACTGGAGGA GAACAAGTGT AAGTAGGCCT TTGTTAGATT CTCCATGGTT 100
CATATCTCAT CTGAGTCTGT TCTCACGCTC CCTCTCTCCC CAGGCTGTGG 150
GGCCCCATCA CCCAGATATT TCCCACAGTT CGGCCTGCTG ACCTAACCAG 200
AGTCATCATG CCTCTTGAGC AAAGAAGTCA GCACTGCAAG CCTGAGGAAG 250
CCTTCAGGCC CAAGAAGAAG ACCTGGGCCT GGTGGGTGCA CAGGCTCTCC 300
AAGCTGAGGA GCAGGAGGCT GCCTTCTTCT CCTCTACTCT GAATGTGGGC 350
ACTCTAGAGG AGTTGCCTGC TGCTGAGTCA CCAAGTCCTC CCCAGAGTCC 400
TCAGGAAGAG TCCTTCTCTC CCACTGCCAT GGATGCCATC TTTGGGAGCC 450
TATCTGATGA GGGCTCTGGC AGCCAAGAAA AGGAGGGGCC AAGTACCTCG 500
CCTGACCTGA TAGACCCTGA GTCCTTTTCC CAAGATATAC TACATGACAA 550
GATAATTGAT TTGGTTCATT TATTCTCCGC AAGTATCGAG TCAAGGGGCT 600
GATCACAAAG GCAGAA 616
ATG CTG GGG AGT GTC ATC AAA AAT TAT GAG GAC TAC TTT CCT 658
GAG ATA TTT AGG GAA GCC TCT GTA TGC ATG CAA CTG CTC TTT 700
GGC ATT GAT GTG AAG GAA GTG GAC CCC ACT AGC CAC TCC TAT 742
GTC CTT GTC ACC TCC CTC AAC CTC TCT TAT GAT GGC ATA CAG 784
TGT AAT GAG CAG AGC ATG CCC AAG TCT GGC CTC CTG ATA ATA 826
GTC CTG GGT GTA ATC TTC ATG GAG GGG AAC TGC ATC CCT GAA 868
GAG GTT ATG TGG GAA GTC CTG AGC ATT ATG GGG GTG TAT GCT 910
GGA AGG GAG CAC TTC CTC TTT GGG GAG CCC AAG AGG CTC CTT 952

ACC CAA AAT TGG GTG CAG GAA AAG TAC CTG GTG TAC CGG CAG 994
GTG CCC GGC ACT GAT CCT GCA TGC TAT GAG TTC CTG TGG GGT 1036
CCA AGG GCC CAC GCT GAG ACC AGC AAG ATG AAA GTT CTT GAG 1078
TAC ATA GCC AAT GCC AAT GGG AGG GAT CC 1107

(2) INFORMATION FOR SEQUENCE ID NO: 24:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: genomic DNA
(ix) FEATURE:
(A) NAME/KEY: smage-1
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

TCTGTCTGCA TATGCCTCCA CTTGTGTGTA GCAGTCTCAA ATGGATCTCT 50
CTCTACAGAC CTCTGTCTGT GTCTGGCACC CTAAGTGGCT TTGCATGGGC 100
ACAGGTTTCT GCCCCTGCAT GGAGCTTAAA TAGATCTTTC TCCACAGGCC 150
TATACCCCTG CATTGTAAGT TTAAGTGGCT TTATGTGGAT ACAGGTCTCT 200
GCCCTTGTAT GCAGGCCTAA GTTTTTCTGT CTGCTTAACC CCTCCAAGTG 250
AAGCTAGTGA AAGATCTAAC CCACTTTTGG AAGTCTGAAA CTAGACTTTT 300
ATGCAGTGGC CTAACAAGTT TTAATTTCTT CCACAGGGTT TGCAGAAAAG 350
AGCTTGATCC ACGAGTTCAG AAGTCCTGGT ATGTTCCTAG AAAG 394
ATG TTC TCC TGG AAA GCT TCA AAA GCC AGG TCT CCA TTA AGT 436
CCA AGG TAT TCT CTA CCT GGT AGT ACA GAG GTA CTT ACA GGT 478
TGT CAT TCT TAT CCT TCC AGA TTC CTG TCT GCC AGC TCT TTT 520
ACT TCA GCC CTG AGC ACA GTC AAC ATG CCT AGG GGT CAA AAG 562
AGT AAG ACC CGC TCC CGT GCA AAA CGA CAG CAG TCA CGC AGG 604
GAG GTT CCA GTA GTT CAG CCC ACT GCA GAG GAA GCA GGG TCT 646
TCT CCT GTT GAC CAG AGT GCT GGG TCC AGC TTC CCT GGT GGT 688
TCT GCT CCT CAG GGT GTG AAA ACC CCT GGA TCT TTT GGT GCA 730
GGT GTA TCC TGC ACA GGC TCT GGT ATA GGT GGT AGA AAT GCT 772
GCT GTC CTG CCT GAT ACA AAA AGT TCA GAT GGC ACC CAG GCA 814
GGG ACT TCC ATT CAG CAC ACA CTG AAA GAT CCT ATC ATG AGG 856

AAG GCT AGT GTG CTG ATA GAA TTC CTG CTA GAT AAA TTT AAG 898
ATG AAA GAA GCA GTT ACA AGG AGT GAA ATG CTG GCA GTA GTT 940
AAC AAG AAG TAT AAG GAG CAA TTC CCT GAG ATC CTC AGG AGA 982
ACT TCT GCA CGC CTA GAA TTA GTC TTT GGT CTT GAG TTG AAG 1024
GAA ATT GAT CCC AGC ACT CAT TCC TAT TTG CTG GTA GGC AAA 1066
CTG GGT CTT TCC ACT GAG GGA AGT TTG AGT AGT AAC TGG GGG 1108
TTG CCT AGG ACA GGT CTC CTA ATG TCT GTC CTA GGT GTG ATC 1150
TTC ATG AAG GGT AAC CGT GCC ACT GAG CAA GAG GTC TGG CAA 1192
TTT CTG CAT GGA GTG GGG GTA TAT GCT GGG AAG AAG CAC TTG 1234
ATC TTT GGC GAG CCT GAG GAG TTT ATA AGA GAT GTA GTG CGG 1276
GAA AAT TAC CTG GAG TAC CGC CAG GTA CCT GGC AGT GAT CCC 1318
CCA AGC TAT GAG TTC CTG TGG GGA CCC AGA GCC CAT GCT GAA 1360
ACA ACC AAG ATG AAA GTC CTG GAA GTT TTA GCT AAA GTC AAT 1402
GGC ACA GTC CCT AGT GCC TTC CCT AAT CTC TAC CAG TTG GCT 1444
CTT AGA GAT CAG GCA GGA GGG GTG CCA AGA AGG AGA GTT CAA 1486
GGC AAG GGT GTT CAT TCC AAG GCC CCA TCC CAA AAG TCC TCT 1528
AAC ATG TAG 1537
TTGAGTCTGT TCTGTTGTGT TTGAAAAACA GTCAGGCTCC TAATCAGTAG 1587
AGAGTTCATA GCCTACCAGA ACCAACATGC ATCCATTCTT GGCCTGTTAT 1637
ACATTAGTAG AATGGAGGCT ATTTTTGTTA CTTTTCAAAT GTTTGTTTAA 1687
CTAAACAGTG CTTTTTGCCA TGCTTCTTGT TAACTGCATA AAGAGGTAAC 1737
TGTCACTTGT CAGATTAGGA CTTGTTTTGT TATTTGCAAC AAACTGGAAA 1787
ACATTATTTT GTTTTTACTA AAACATTGTG TAACATTGCA TTGGAGAAGG 1837
GATTGTCATG GCAATGTGAT ATCATACAGT GGTGAAACAA CAGTGAAGTG 1887
GGAAAGTTTA TATTGTTAAT TTTGAAAATT TTATGAGTGT GATTGCTGTA 1937
TACTTTTTTC TTTTTTGTAT AATGCTAAGT GAAATAAAGT TGGATTTGAT 1987
GACTTTACTC AAATTCATTA GAAAGTAAAT CGTAAAACTC TATTACTTTA 2037
TTATTTTCTT CAATTATGAA TTAAGCATTG GTTATCTGGA AGTTTCTCCA 2087
GTAGCACAGG ATCTAGTATG AAATGTATCT AGTATAGGCA CTGACAGTGA 2137
GTTATCAGAG TCT 2150

(2) INFORMATION FOR SEQUENCE ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2099 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:

(A) NAME/KEY: smage-II

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

ACCTTATTGG GTCTGTCTGC ATATGCCTCC ACTTGTGTGT AGCAGTCTCA 50
AATGGATCTC TCTCTACAGA CCTCTGTCTG TGTCTGGCAC CCTAAGTGGC 100
TTTGCATGGG CACAGGTTTC TGCCCCTGCA TGGAGCTTAA ATAGATCTTT 150
CTCCACAGGC CTATACCCCT GCATTGTAAG TTTAAGTGGC TTTATGTGGA 200
TACAGGTCTC TGCCCTTGTA TGCAGGCCTA AGTTTTTCTG TCTGCTTAGC 250
CCCTCCAAGT GAAGCTAGTG AAAGATCTAA CCCACTTTTG GAAGTCTGAA 300
ACTAGACTTT TATGCAGTGG CCTAACAAGT TTTAATTTCT TCCACAGGGT 350
TTGCAGAAAA GAGCTTGATC CACGAGTTCG GAAGTCCTGG TATGTTCCTA 400
GAAAGATGTT CTCCTGGAAA GCTTCAAAAG CCAGGTCTCC ATTAAGTCCA 450
AGGTATTCTC TACCTGGTAG TACAGAGGTA CTTACAGGTT GTCATTCTTA 500
TCTTTCCAGA TTCCTGTCTG CCAGCTCTTT TACTTCAGCC CTGAGCACAG 550
TCAACATGCC TAGGGGTCAA AAGAGTAAGA CCCGCTCCCG TGCAAAACGA 600
CAGCAGTCAC GCAGGGAGGT TCCAGTAGTT CAGCCCACTG CAGAGGAAGC 650
AGGGTCTTCT CCTGTTGACC AGAGTGCTGG GTCCAGCTTC CCTGGTGGTT 700
CTGCTCCTCA GGGTGTGAAA ACCCCTGGAT CTTTTGGTGC AGGTGTATCC 750
TGCACAGGCT CTGGTATAGG TGGTAGAAAT GCTGCTGTCC TGCCTGATAC 800
AAAAAGTTCA GATGGCACCC AGGCAGGGAC TTCCATTCAG CACACACTGA 850
AAGATCCTAT CATGAGGAAG GCTAGTGTGC TGATAGAATT CCTGCTAGAT 900
AAGTTTAAGA TGAAAGAAGC AGTTACAAGG AGTGAAATGC TGGCAGTAGT 950
TAACAAGAAG TATAAGGAGC AATTCCCTGA GATCCTCAGG AGAACTTCTG 1000
CACGCCTAGA ATTAGTCTTT GGTCTTGAGT TGAAGGAAAT TGATCCCAGC 1050
ACTCATTCCT ATTTGCTGGT AGGCAAACTG GGTCTTTCCA CTGAGGGAAG 1100
TTTGAGTAGT AACTGGGGGT TGCCTAGGAC AGGTCTCCTA ATGTCTGTCC 1150
TAGGTGTGAT CTTCATGAAG GGTAACCGTG CCACTGAGCA AGAGGTCTGG 1200
CAATTTCTGC ATGGAGTGGG GGTATATGCT GGGAAGAAGC ACTTGATCTT 1250
TGGCGAGCCT GAGGAGTTTA TAAGAGATGT AGTGCGGGAA AATTACCTGG 1300
AGTACCGCCA GGTACCTGGC AGTGATCCCC CAAGCTATGA GTTCCTGTGG 1350
GGACCCAGAG CCCATGCTGA AACAACCAAG ATGAAAGTCC TGGAAGTTTT 1400
AGCTAAAGTC AATGGCACAG TCCCTAGTGC CTTCCCTAAT CTCTACCAGT 1450
TGGCTCTTAG AGATCAGGCA GGAGGGGTGC CAAGAAGGAG AGTTCAAGGC 1500

AAGGGTGTTC ATTCCAAGGC CCCATCCCAA AAGTCCTCTA ACATGTAGTT 1550
GAGTCTGTTC TGTTGTGTTT GAAAAACAGT CAGGCTCCTA ATCAGTAGAG 1600
AGTTCATAGC CTACCAGAAC CAACATGCAT CCATTCTTGG CCTGTTATAC 1650
ATTAGTAGAA TGGAGGCTAT TTTTGTTACT TTTCAAATGT TTGTTTAACT 1700
AAACAGTGCT TTTTGCCATG CTTCTTGTTA ACTGCATAAA GAGGTAACTG 1750
TCACTTGTCA GATTAGGACT TGTTTTGTTA TTTGCAACAA ACTGGAAAAC 1800
ATTATTTTGT TTTTACTAAA ACATTGTGTA ACATTGCATT GGAGAAGGGA 1850
TTGTCATGGC AATGTGATAT CATACAGTGG TGAAACAACA GTGAAGTGGG 1900
AAAGTTTATA TTGTTAGTTT TGAAAATTTT ATGAGTGTGA TTGCTGTATA 1950
CTTTTTTCTT TTTTGTATAA TGCTAAGTGA AATAAAGTTG GATTTGATGA 2000
CTTTACTCAA ATTCATTAGA AAGTAAATCA TAAAACTCTA TTACTTTATT 2050
ATTTTCTTCA ATTATTAATT AAGCATTGGT TATCTGGAAG TTTCTCCAG 2099

(2) INFORMATION FOR SEQUENCE ID NO: 26:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Glu Ala Asp Pro Thr Gly His Ser Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 27:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Glu Val Val Pro Ile Ser His Leu Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Glu Val Val Arg Ile Gly His Leu Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Glu Val Asp Pro Ile Gly His Leu Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Gly Val Asp Pro Ala Ser Asn Thr Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Glu Val Asp Pro Thr Ser Asn Thr Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

Glu Ala Asp Pro Thr Ser Asn Thr Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Glu Ala Asp Pro Thr Ser Asn Thr Tyr
5

(2) INFORMATION FOR SEQUENCE ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Glu Val Asp Pro Ile Gly His Val Tyr
5

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: single

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

Tyr Met Asn Gly Thr Met Ser Gln Val
5

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Met Leu Leu Ala Val Leu Tyr Cys Leu Leu
5 10

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

Met Leu Leu Ala Val Leu Tyr Cys Leu
5

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

Leu Leu Ala Val Leu Tyr Cys Leu Leu
5

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

Ser Glu Ile Trp Arg Asp Ile Asp Phe Ala His Glu Ala
5 10

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

Ser Glu Ile Trp Arg Asp Ile Asp Phe Ala
5 10

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

Ser Glu Ile Trp Arg Asp Ile Asp Phe
5

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

Glu His Ser Ala Tyr Gly Glu Pro Arg Lys Leu Leu Thr Gln Asp Leu
5 10 15

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

Glu His Ser Ala Tyr Gly Glu Pro Arg Lys Leu Leu
5 10

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

Ser Ala Tyr Gly Glu Pro Arg Lys Leu
5

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1032 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

CGCCAATTTA GGGTCTCCGG TATCTCCCGC TGAGCTGCTC TGTTCCCGGC TTAGAGGACC 60
AGGAGAAGGG GGAGCTGGAG GCTGGAGCCT GTAACACCGT GGCTCGTCTC ACTCTGGATG 120
GTGGTGGCAA CAGAGATGGC AGCGCAGCTG GAGTGTTAGG AGGGCGGCCT GAGCGGTAGG 180
AGTGGGGCTG GAGCAGTAAG ATGGCGGCCA GAGCGGTTTT TCTGGCATTG TCTGCCCAGC 240
TGCTCCAAGC CAGGCTGATG AAGGAGGAGT CCCCTGTGGT GAGCTGGAGG TTGGAGCCTG 300

AAGACGGCAC AGCTCTGTGC TTCATCTTCT GAGGTTGTGG CAGCCACGGT GATGGAGACG 360
GCAGCTCAAC AGGAGCAATA GGAGGAGATG GAGTTTCACT GTGTCAGCCA GGATGGTCTC 420
GATCTCCTGA CCTCGTGATC CGCCCGCCTT GGCCTTCCAA AGTGCCGAGA TTACAGCGAT 480
GTGCATTTTG TAAGCACTTT GGAGCCACTA TCAAATGCTG TGAAGAGAAA TGTACCCAGA 540
TGTATCATTA TCCTTGTGCT GCAGGAGCCG GCTCCTTTCA GGATTTCAGT CACATCTTCC 600
TGCTTTGTCC AGAACACATT GACCAAGCTC CTGAAAGATG TAAGTTTACT ACGCATAGAC 660
TTTTAAACTT CAACCAATGT ATTTACTGAA AATAACAAAT GTTGTAAATT CCCTGAGTGT 720
TATTCTACTT GTATTAAAAG GTAATAATAC ATAATCATTA AAATCTGAGG GATCATTGCC 780
AGAGATTGTT GGGGAGGGAA ATGTTATCAA CGGTTTCATT GAAATTAAAT GTTATCAACG 840
GTTTCATTGA AATTAAATCC AAAAAGTTAT TTCCTCAGAA AAATCAAATA AAGTTTGCAT 900
GTTTTTTATT CTTAAAACAT TTTAAAAACC ACTGTAGAAT GATGTAAATA GGGACTGTGC 960
AGTATTTCTG ACATATACTA TAAAATTATT AAAAAGTCAA TCAGTATTCA ACATCTTTTA 1020
CACTAAAAAG CC 1032

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

Met Ala Ala Arg Ala Val Phe Leu Ala Leu Ser Ala Gln Leu Leu Gln
5 10 15

Ala Arg Leu Met Lys Glu
20

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

Met Ala Ala Arg Ala Val Phe Leu Ala Leu Ser Ala Gln Leu Leu Gln
5 10 15

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

Ala Ala Arg Ala Val Phe Leu Ala Leu
5

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 354 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

ATG CGA AGA GAA GAT GCT CAC TTC ATC TAT GGT TAC CCC AAG AAG GGG 48
Met Pro Arg Glu Asp Ala His Phe Ile Tyr Gly Tyr Pro Lys Lys Gly
5 10 15

GAC GGC CAC TCT TAC ACC ACG GCT GAA GAG GCC GCT GGG ATC GGC ATC 96
His Gly His Ser Tyr Thr Thr Ala Glu Glu Ala Ala Gly Ile Gly Ile
20 25 30

CTG ACA GTG ATC CTG GGA GTC TTA CTG CTC ATC GGC TGT TGG TAT TGT 144
Leu Thr Val Ile Leu Gly Val Leu Leu leu Ile Gly Cys Trp Tyr Cys
35 40 45

AGA AGA CGA AAT GGA TAC AGA GCC TTG ATG GAT AAA AGT CTT CAT GTT 192
Arg Arg Arg Asn Gly Tyr Arg Ala Leu Met Asp Lys Ser Leu His Val
50 55 60

GGC ACT CAA TGT GCC TTA ACA AGA AGA TGC CCA CAA GAA GGG TTT GAT 240

Gly Thr Gln Cys Ala Leu Thr Arg Arg Cys Pro Gln Glu Gly Phe Asp
65 70 75 80

CAT CGG GAC AGC AAA GTG TCT CTT CAA GAG AAA AAC TGT GAA CCT GTG 288
His Arg Asp Ser Lys Val Ser Leu Gln Glu Lys Asn Cys Glu Pro Val
85 90 95

GTT CCC AAT GCT CCA CCT GCT TAT GAG AAA CTC TCT GCA GAA CAG TCA 336
Val Pro Asn Ala Pro Pro Ala Tyr Glu Lys Leu Ser Ala Glu Gln Ser
100 105 110

CCA CCA CCT TAT TCA CCT 354
Pro Pro Pro Tyr Ser Pro
115

(2) INFORMATION FOR SEQ ID NO: 50:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 676 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

TCTTCATACA CGCGGCCAGC CAGCAGACAG AGGACTCTCA TTAAGGAAGG TGTCCTGTGC 60
CCTGACCCTA CAAGATGCCA AGAGAAGATG CTCACTTCAT CTATGGTTAC CCCAAGAAGG 120
GGCACGGCCA CTCTTACACC ACGGCTGAAC AGGCCGCTGG GATCGGCATC CTGACAGTGA 180
TCCTGGGAGT CTTACTGCTC ATCGGCTGTT GGTATTGTAG AAGACGAAAT GGATACAGAG 240
CCTTGATGGA TAAAAGTCTT CATGTTGGCA CTCAATGTGC CTTAACAAGA AGATGCCCAC 300
AAGAAGGGTT TGATCATCGG GACAGCAAAG TGTCTCTTCA AGAGAAAAAC TGTGAACCTG 360
TGGTTCCCAA TGCTGCAGGT GCTTATGAGA AACTCTCTGC AGAACAGTCA GGACCACCTT 420
ATTCACCTTA AGAGCCAGCG AGACACCTGA GACATGGCTG AAATTATTTC TCTCACACTT 480
TTGCTTGAAT TTAATACAGA CATCTAATGT TCTCCTTTGG AATCCTGTAG GAAAAATGCA 540
AGCCATCTCT AATAATAAGT CAGTGTTAAA ATTTTAGTAG GTCCGCTAGC AGTACTAATC 600
ATGTGAGGAA ATGATGAGAA ATATTAAATT GGGAAAACTC CATCAATAAA TGTTGCAAAT 660
GCATAGTAAA AAAAAA 676

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13585 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: at positions 9421-9456, the "Ns" refer to an unsequenced portion of from 4.7 to 5.3 kilobases

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

CCGTCAGAAA TCTAAACCCG TGACTATCAT GGGACTCAAA ACCAGCCCAA AAAATAAGTC 60
AAAACGATTA AGAGCCAGAG AAGCAGTCTT CATACACGCG GCCAGCCAGC AGACAGAGGA 120
CTCTCATTAA GGAAGGTAAG AGCGTTGCCT TCTCGCCATA ATCATAGTCC TCTTCTCCCA 180
GAATAGGATT TGGGAAATTC TGGCTAAGTC CTCTGCCTAC CCTCATTGCC CCGCTGATGT 240
GTGACATCAA CAGAATTTCT CCGCAACGTT TGTCAGTCTC CAACCTCAGA GGGCTCACAA 300
AGCCTCCTCC TGAATCCTCT CTCAGTCCTC CAACACTACC AAGAAGAAAA GCAATTATTC 360
AGGATGGCAT CTTGCTGGGG AGAAGCAGCC TCCCTGAGGT AGATGTGTTC TCCTGTCACT 420
TAAAGAACCA CTTCTCCTGG TCTGAGTAGT AAGAGGCGCA TTTGCTGTTG CTGCACCATT 480
TGCCAAGGCT CTGAGTTTGA GGTATGGGAT GTATTAAAAC AATTTAATGA AGAATTAAGA 540
TTCCATTCTG TCATTTTGAA CACAGGGTTC AGTCCTATAT TATTCACTTG AGAGGACTGG 600
TGAGTTTGAC TTTCATTTCT TTTTTACAAC TGGGAAGGGC AAATTACACA TAAAATGTCC 660
CAGTGGAAAG GGGTCATGTG TCGAAATCCC CACTCTTCTG TCTCACCTCT CCCTGTTGTT 720
TTAAACTGGG GCTCATTAAT ATAATTCTAT GGGGATCACA CCTTTGAAAT TCATGAGGAC 780
AGTAAGAGAG CAGAAAAATA CACAATAATA AGGAAAGGAG CTTCCATTAT TGGTTTTTAA 840
TGAGCGTACT TGAATTACGG CCACTGCaGT TTATGGATAT TTTTTGTTGT TCATTTGTAT 900
GTGTTATAGT TAGAAAAAAA AAGAATCCTA GCCAAGGGAC TTGAACCAGA GAGAAGCAGA 960
AATTGACTTA AGTAGGAAGG GAAACACATT ATTAGATAAA GTCAGGTCCT GGGCTTCCTC 1020
GGCTTGTTTT GGGTGGAGTG CCTGGGGACA GGCTGAAGCC CCTGTGTGGG GTGGTTTCCT 1080
TTGCTGAAAA GCTGGGCTGG AAGATGTTGT GCTCAGTGCT CAACCTCATG CACCCTCGCG 1140
AGGCACAGGC AACGGGTGCT CTGGGAAACA CACGTTATGT ATCATAGCCT CTGTTTGTCT 1200
GTGGGATTGA TATCCAATAA TAACTTTGGA GAAAAATAAC TCCTCTTATT TTGTTAGCCA 1260
CAGCCCTGGG CCAGGGAAGG TGGAGAATCA GTGAAAATGC ATTTTGTTTG TTTCTCTAGA 1320

AGTTTATGGT GCAGAGTCAA ATTGAAGGCA AATGAGGAAT ATTTTTTCAT TAAATAATAA 1380
CTCAACTTGC AAGTCTTTTT TGCTTTTGTT TGTAGTTTCT TCTTTGAACT TAATTTTCAG 1440
TTAGTAGGAG GGGTTAGAAA CCTGAGCTAT TGCTAAAGCC CTTGATATGA ATGAAAGAAG 1500
CAGGTGCAAA TCCCCTCACA GAGAGAAACC AAAGGGTCCT GGCTATGGAT ATTGGTCACC 1560
TAGTCAGGAT GCTGTTGTGG GTCTTTATGA GATGATGAAT AGGGTGGCTT TGGATGCATT 1620
AATGATATTT ACATGCTCCT TCTGTTAGTG TCCTGTGCCC TGACCCTACA AGATGCCAAG 1680
AGAAGATGCT CACTTCATCT ATGGTTACCC CAAGAAGGGG CACGGCCACT CTTACACCAC 1740
GGCTGAAGAG TAAGTTCAAA ACCAGACCCA GCAGGGCTTC CAGTTTGCCG TTTGCTGACA 1800
CAGCCTGCTG ACTTCCACCA GTACATGCCT GCTCGTAAAT CTCCCTAGTG TTTATCTCCC 1860
CAGACAGTAA CATCCCTGGC AACAAGGGGA GGAGATTCTG TGCTTCTATA AGGGGCTCAG 1920
TCAAGCTTCT CTGAGGCCAA ACAGGCAGGA AGATGGGAAT GGTATAAGGT TGGATCTTGC 1980
CATTTTTGGG TGCACTTTTG ACTATTGGGT CTTATCTGTA GGTTCCCAAG TGGAAAAACA 2040
TCTGTTCAGG ATCACAATGC CTCTCTCCTC AATCCTTGTT CTGTCTCCTC CACTCAAATT 2100
CCTGAAGGTG GTTTGCAGAC AGAATAAAAG TGAGTTGCCA AGGAGCCAGT AAGGATGACG 2160
GGCAGGTGTG TGTGACTCAG CCCACAGCCA GACTCGAGAG GAAGATGGAG GTCACAGCCT 2220
TTGCAGTATA ACTTTATCCT AAGGAAAGAC ATTGGGTTTT ATGAGTGAAT TAAAAATAAG 2280
TATTTATATG ATTAAGCATT TCTAAATGCT AAGCATTGTA TACTGGCGTG AGACACTGTT 2340
TTTATCTTTG AAAAAACTCA CAACTTAGTG GGAGAGTTAG GCATGAGATT AATTTCAGCA 2400
AATGTAAGTG CGGTAATGAA AACCCAGAGG CTGCAGGGAC ATACTCTGTA TGTGCTGGGA 2460
GTGGGAAAGG GACATACTCT GTACGTGCTG GGTGGCAGGG GCAGGGGAGG CCCCACCCTC 2520
TGCGTGGGAC TGTAACAGGA CAACACCCTC TTATGTGGTC TGTCCAGAAC TCCCTGTGAA 2580
CCTGCTCTTT CTTTGGAAAG AGCTGTTGAA CAATCTTTGT TAACAGTCAA CCGCAGGACC 2640
AGCAAGATGT AAAGCCCAAC AAAGGCACTG AGGAAGAGTT CAGGAAGACA GCATTTCCTC 2700
AGAAGACCCT GGTATAGGAT CCTCTAATAT CCCTGGCCAA TTGGAGATGA GGGCGGCGGT 2760
ATCCTCTCAG AAAATGTCCT GACAGCAAAA ACATACTCTT TGAGGGAGGG GAGCCCATTG 2820
CCCGTGCTAT TAGTTAGGGT ATCGTTTCAG CTTGTGTATA ATCACTCAAC AGACTCTTTA 2880
AAATATACTT TTATGTCTCG TGTAAAAATT CAAGAGTAAA GAGTTCAAGG CCTGTTCGTT 2940
TTCTTCTTGC TGGTTACTCC CTTGGGATCG TCACTTTTGT CCCCATGGCT GAAGATGTTG 3000
TGCCATCACC TCCACATCTT GCCAACAGAA AGCAGGAGGT GAAGGAGAGG CTAGGACCAT 3060
TCCTTTCAAG GGGCACACGT CACTTCTGCT TATTGCTCCA CCCCCGCCCC CCGCCCCGTG 3120
GCACCCACCC TGGTGGTATC ATTCTTGCTG TGTTGTAAAT GAAGAAAGGT TTAGAGAAAT 3180
TAGGAAATGT GTGGCCAGAC ATGGTGGCGC TGGGATTTAA ATCCAGGTCT GTTTGCCTCC 3240
AGAGTCCATG CTCTTAAGTG TTATGCTGCA GGCCAGCAGA GGCAAATATT TGCACAATCC 3300
CATCCGACGA GAGGCTAGGG CAGAGGTCAG TATCTCTCAG TGTGAAGCTG GAGGCTGATG 3360
CTAGTCAGCT CAGTAGGCCG AAAGTGGAGT TGTCCTTTGC CATGTAGGGC CATCATGCCC 3420
AGCTGGGGAA CCTCATAGCC AGGTGTACCC ACAACCTGAA CAAGGTAACT TTCAGGGTCT 3480
AGTCAGGAAG AAACCAACTA GATGGTTCAA CATAGAGACT TTAATATAAG AAGCTGGTTA 3540

AACAGGCATG GGACTGAGAC TGAGGAGGCA AAGAAGGCAT CGGGGCAACC AAGGCTGTAC 3600
CCACAGAATG CTGCTTCTAC CCCCGTGTCT GGGGTAACAA ACGGAAGGGT GAGGCCATCA 3660
GGACCTAGAG TTGGGAGGAG GGACGCCACA GAAATGGGAC CCAGATCTCT AAGGAGAGAT 3720
TTTTGTTTGG CTGGTTCTGG TGTCTCAAGA GCTTAGAAGT GAGGGGCATG AATCAAATAC 3780
TCAGGCCTCT GAGGTCAGCC AGTGCTCTGC TGGGGAGGGG CATAATGAAG CTGGCTCTGA 3840
CAATGCCGGA AAACGAGCTG GTGCTTGGCA TATACAGACA ATGTGAGCAT TGCTGGGGTG 3900
ATCCTGACAG GAGCCAGAAG CACACTGGAA GGAGCTGCTC CTTCTTGATG CCCCAGGTTT 3960
GTAGGCACCC TCTAGAGTAC TCTAATGGGA GCCAGTGGGC AAAGGAGAAG TGGCATTTGC 4020
AGAGTCCAGT CCCAGCATCA CAGAGCAGAG CATAGAAAGG TAGGTTTGGA GAAGAGGGAC 4080
AATGGCTTAA TAAAGGGCAA AGGGGGTTAT GACCACTATC ATGTGAAGGA ACCCCTTGAC 4140
TGAAGGCACA AGCTTTCTGT GTCTTGCAAC CTGAATGACG TGCATAAGCA GGGTCAGGTG 4200
GGTTATCTGA CATTTTCCTT GAGAACAAGA GGGAGCCTCT GGATTCCAGC ACAAAAGAAA 4260
AATACCCACT CAACCCGTAT GCGTGGGAGC TATCCTTTAA AGAGAAAGTA ATTCCTTTTG 4320
ACATTTTGCT GTCTGTAGAA GGGTCAGATG GCCAAAGCTT CCAGCACAAT GAAACACTTA 4380
ACTTCAGTCT GTGAGTGTAG GAACCCCTGA ATACATGGAA CATCATCATC TTGTGCAGGT 4440
ACTGAAGGAG ATCGGTCCAG AAAATAAGTA ACTGCACATG GCCACCAATG TCAAAAGTCA 4500
TTCCTCTCAT GAAAAGTCCC TGCCCCCATT GCTGTTTGTT TAAATAGGTG GGATGGAGGT 4560
AGGGGAATGG GGCCATCTTC TTTTTTTTTT TTTAATTTTT TTGCATAAAA TCCAGATCCT 4620
GCACAATGGG GCAATCTTCA TTAAAACAAT GCATCCCTAA GATCTGAGAA TATTTATCCT 4680
TCTCACAATT GTGCCAGCAG GTGGAATGAA GAAGAATGAT GCAAAATAAG TTCCCACATC 4740
CAGCCAAGAA GGACTACATA CCTGCTTTGG GTATTATGTA TCCCTTTGAA ACCTCAGTGG 4800
AGAGCAGTTC TCACAGTTGG GTGGACACAA GTCATCCATG GAACTTGTTA AAATGCAGAT 4860
TTCTAGGTGC TGCCACCTAA GAGGCTGATT GGGTAGGCCA GGGGTGGAGT CCTATGATCT 4920
GCACCTTAAC GTGCATCTCA GGTGATTCTG CTGCAGGTGG TATTTGGAAG ACACTCTGAG 4980
GCGCCCTGCC AAGCTGGGCA GTGGGTTCTT CCAATGTGTC AGGCATACCC TGGTGCTTTT 5040
CGCTCTCAGT CACTTGGGCA TGTTGTGAGT ACCACGTGAC CATGCATAAA GTGCTGTAAC 5100
AGAGCTCTGT CTGTGTCAAG ATATTCAAGT GGACGCCACA GGGTAAAATG AGAGCACAGG 5160
CATGTTGGGA GTTGAATCAG CTGCCTTCAG TCACGAGAAC ACACTGAACA CTCCTTGTGA 5220
CAGCTTCAGT TCAGGAAAGA GTGACTCTGC AGGAAAAGCA CTGGCCTGGG AGACCTGGAT 5280
CTGGCCCAAA TTCTGGTGCT CACTTGCTTG GTCTCCCGTT CCAGTTGCTG TGAATGTTGG 5340
TTCTGCCACT TGCTGGTTGT GCAGCCCTGG GCACTTGACC AGCATAATGT CAGCTGTAAA 5400
ATGAACATCA TTCCTAACTC CGAGGACTGT GGTTAGGATG AAATAAAAGC ATATATGTGG 5460
GGGTGCCTAG CCCAGTGCCT GGCACAAATT GGTGCTCAAT GAATGGTAGT CACTATGGTT 5520
ATGGTAATGT TGATGAATCT TCATAGGTCT CAGCTTCCTG ATCTATAAAG CGGGTGGACT 5580
GACCTACATA AGTCAGAGTT TCCATCTAGC ACTGTCATCC CATGGTTCGC TCTATCCTGT 5640
TTGGAGACGG ACAGGATAAG CTTGATGTCT CCTCAGCCTT GAGACAGAAG TTGTCCAGTA 5700

GATGGTACTG AGCAAAAGTC TCTCCAGCAG AAGCCTTAGT TAAACCTTGC TTCTCCTGTA 5760
GCTGCTCAGT CTCTTGTAAG TCACTCAGCT CTGCAGAAAC TTTCTTAGCG AGTTGACAAC 5820
CACAGATAAC AGAGTCAGTT CTGTCGATTT TGATCATGCT GTGATCAGGC AGATGTTAGC 5880
TAATTGATGA TGCTTGCCCG GAGTGAACAG CTCCAGGCCC TGTTTCCAGG GTCTTTGTGG 5940
TAACTTTGTG GTAACTGTAA TGCTTCCCAG GGGTCACTGA ACACAGGGCC CAAGAGGCTG 6000
GTGTAGACCC CCAGATTGGC ACCCTGCTGC TTAGACAAGA TCCTTCTCAA TAAGTAATGC 6060
CATAGCTTTG CTGTAGGTTC AGCCCAGACA CTTCTCCCTA GGGCTGCAAG GAGCAAAGCG 6120
GGGAGTTTAG GGAAGGGAGG GCACGAACAT AATTGAGACG GATTCAGGTT CAAATCCAGC 6180
CTCTGTTTTG TGCTAGCTCT GTATGATCAC CAGCGAGTCA TGTATCCTCT GCCTTTTATT 6240
TCCTCTTCTG TGAAAATAGG GGATGATAAA TTGTGTCTAC CCTCCAGTGT TGATGTGAGA 6300
ATTGAATAAG CTAATGAATG TTTAGCACAG CACCTGGCTT TTAGTAGATG AGTCAGTGTT 6360
AATTTCTATT TTCTCTTTGT GGGCTGAGTT GGAGAAAATG TTTTAAAACA GCCTGATGAG 6420
AAGAAAAGAT AATTTAGCCC CAATAAATAC ATTGTCCACA TAAAGACAGT TACTATGGCA 6480
CTTCTCATAC CTGGAACTTG GGTGCCTGGG CCATGCAATT AGCAGAGTTC CTGTGGGCAC 6540
ACACTTGAGA GGCTCCTAAA GACCTGGGTT AGATCCAGGT GCTGGAGGCC TGGTGGGGTG 6600
CCAGTGTGGG AGGTGGGAAA CTACTTGGAC ACTGGGAGAT GCTGCTCTGG GTCGTCAAAG 6660
TCCATATGAA GAGGAAGACT GATTTATGCT TCATCATAAT GTAGAACAAT GTTTCAATGA 6720
CAAAGTGGAT TTGTCTATCT CTTGGGCCAG GCCGCTGGGA TCGGCATCCT GACAGTGATC 6780
CTGGGAGTCT TACTGCTCAT CGGCTGTTGG TATTGTAGAA GACGAAATGG ATACAGAGCC 6840
TTGATGGTTG GTAAAGTTCC CACTGCTGAA ATCCCTCCAA GTCCAGGGCC CTCTTTCCAG 6900
TTCTTTCCTC TGAATCTCTG GAGAGTCAGA TAATTGCCTC ATTATAACCT TCAGCTCTGA 6960
TTCCGGCTTC TGATGCCTCT TTTGCTACAT TGTACTTTGG CAACTCTACC TTTGCCTCTG 7020
CTCAGGCATG AACCTCAACC AGGAACTTGC CCTGTGTCTT AGTCTGTGAT TATAACATAA 7080
TACGAGAGAC TGTAATTTAT AAATAAATGA AATTCATTTG GTTTACAGTT GGGAGGCTGG 7140
GAACTCCAAG ATCTAGGGGC CACACCTGGT GAGGACTTCT TGCTGTGTCA TATCATAGTG 7200
GAAGGCATCA CATGGGCAAG GGAGTGAGAG AGCAAGAGGG AGCTGAACTC ATTTTTTTTT 7260
TTTCTTGAAA CAGGAAATCC TGGGATGGAG CGCAGTGGTG ATCATGAGTC ACTGTAGCCT 7320
TGACCTCCTG GGCTCAAGCC ATCCTCCTGT CTCAGCCTCC AGAGTAGCTG GGACCACAGG 7380
CACGTGCCAC CACACCGGCT AATTAAAAAA AAACTTTTTT TTGTAGAGAC GAGGTCCCAC 7440
TATGTTGCCC TAGGCTGGTC TCAAACTCCT GGGCTAAAGT GATCCTGCCT CGGCCTCCCA 7500
AAGTGTTGGG ACTACAAGTG TGAAACACTC CACATATGGC CCAAACTCAC TTTTATAACC 7560
AACCTACTTT TGCAATAACA AACACACTCC TGCAATAACA CAATTAATCC ATTCGATGAG 7620
GACAGAGCCC TTGTAACTTA ATCGACCTCT TAAAAGTCCT GCCTGTTACC ATTGTTGCAT 7680
TGGGGATTAG GTTTCCAATA CACGAATTTT GGGGGACACA TTCAAACTAT AGCACCTGTC 7740
TCTTTGGTTC TACTCATAGC AGACTTGGGT ACCTGGATGT TGTGTGTAGC TAAGCACTGA 7800
CGGTTTATAG GGCACAGGGG AAGGGGTTTG AGGTTCCCTT ATAGCAAACA GGAGTATATT 7860
AGACACCTCA GGTTTTACCA CTTCTGGGAA TTCTTGCTGG TTCTGTTACT CCACTTTGTG 7920

ACCTGCTCTT CCTACTTTTC TTCTTCACCC CTTTCCTCAC TGGTTACCTG TGAATTCCAA 7980
GTTCTTCTGA CTCTACACTA AGCATCCCAG GATATCATCA GTGCGATGAG GAAACCATCC 8040
TTCCTGCATC AGCACAAAGG GTCACTTGTG TGTTTTTTAA CAGGCTGCAT CCTTCTTAGA 8100
TGGCCAAAGG TTTTAATAGT ATTTTTTTCT TCTTTACCCA AATATGCAGG AAGCTAACAC 8160
AATTACACAA TCCAATCTTC TGGTACCAGT ATCCTCCATG AATGGGAAAC ATCAACTGAG 8220
TTTATAAGCT ATAAAAATTA CAGGTTTCAG CAATCTTGCT TAAAGCCAGG TAGCACTTCA 8280
GCACTTCAGC ACCCGAAGCA TTCTCCATAG ATCTCGCTGT CTCTCTTTCT TGTTATTACA 8340
GATCTGAAAG CTTTTCAGGT TGATGCATAA TGGAAAAAAA GTATCTTTCC AAAAGATGTT 8400
GGAAAGTCCC ATTCTCATTC AGCAAGCACT TCATTTAGAG GAAAAGGTCC TGTGAAAGAG 8460
AGGAGGGTTG GTGTGGGGTG GGGATTGAAG CTTGGCAAGC TGATAAGGAG AAGGTGAGAG 8520
ATACAACTCT GGATTCTTTC CCTCTTTGCC AAGAAACTTG GGCAGTCTCA TGTCTCATGT 8580
CTCCTGTTCC CCAATGTCTT TCCAGAGCAT AAATACAAAT ACAAACCATC AAAGGCAAGT 8640
CAAGTCTGGG GGCTGACACA CCCACCGAGC ATAGCCCTCT AGTGTGCTGA CATCTAGTGG 8700
GAAGGAGGAG GAGTTGATGA ATCTGAACAA GACTCCAATA TTGGAGGAAA TACTTGAGGA 8760
AAGCCTTGGG TTAGAAAGTT AGGGATAGAA TTCCTGCTCA TACGGCTGTC CACAACAGGT 8820
TAGTAGGGGA GGACTTTAAT CTCTGCCATA GAACTCCATT TGTAACTCTA GCATGGGGTT 8880
ATGACATTGC CTTGTAATTG GCTATTTACT TTTTGCCTCT TCGACCCCTC CGCTTTCCCC 8940
TATGTATGAA CCACAACAGA GAATATTTCT AACTCATCTT CATATCTCCA GTGCCTAGCA 9000
CAGTGCCTGG TACATGGTAG TCACTCAATT GTGTTGCATT AGGACTTGGT CCCATTGTCT 9060
GCCATTGAGT TGCTTGGAGA CTAGAATTCA ACTTCTCCAA GATTCACTAG CTCTATTTTA 9120
CACCCAGACA TGTTGGAAAT CTGTGATGTA ACACAATGTA TATCCATTTT TATTTAATAC 9180
ATATTTTCTT CTATATTTTG ATTTCATTAT ATATTTGTAT ATCAAAAACA AAATGTTTAG 9240
TCTTTCAAGA AGTAAAGCTA TACAAACTCA ATATGTTGGT ACTCATTTCC TAACTATAAT 9300
TATTAGTTTG ATCCTATTGA ACACAAATGC AGTAATTTTT CTTTTCTGCT TCAATGCTCT 9360
CATCTTAAAT TCATTTAATT GAAAAATAAC AGAGAGTCTT AATGTCATGT GCTCAGACAC 9420
TNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNGAGT GCAGTGGTAT GATCTGGGCT 9480
CACTGCAAGC TCTGCCTCCT GGGTTCATGC CATTCTCCTG ACTCAGCCTC CCAAGTAGCT 9540
GGGACTACAG GTGCCCGCTA CCATGCCTGG CTAATTTTTT GTATTTATAG TAGAGATGTC 9600
ATTTCACTGT GTTAGCCAGG ATGGTCTCAA TCTCCTGACC TCGTGATCCA CACGCCTTGA 9660
CCTCCCAAAG TGCTGCGATT ACAGGTGTGA GCCACCGTGC CCGGCCTTAT ACTTCTTTTT 9720
TACTTTTTTT CAGTGGTTTC CCTAGAGTTT GCAACATACA TTTACAACTA ATTCAAATCC 9780
ACTTTCAAAT AACACTATAC CATTTCATAG GCATTATGAG TATCTTAAAA TAATCCTAAT 9840
TCCTTCCTCC TGTAAACTAA AAACAAAATC CTAAATCCTC CAAACAACTG AATGGACCCC 9900
CTCTTCACCA AGGGGACCCC AGGGAAACCT GAAAAACTGA GTGTTGGCCA TGACGGGAAG 9960
GGAGGTGAGA GATGCTCATT ATACTCCCTC CCTTTTAGAG TTTTAGGTAC AACTGACCAG 10020

CATTAATTTT AAAATAGAGA TTACAGGACT GACAGAATGA ACTCTTTGTG GCAATATCAA 10080
ATTAGGAACA AGACAATGCA AGGAAAGGGT TAAATCATGC CCTTCAAACC ATAAAAAAAT 10140
TTTTTTTTAA TTAACCCCAT ATAATGTGGT ATACTTTCCA AACTGACTCT GGTATAGCAT 10200
CACATGACAG ATTGCAGACT CCCTTACCTT AAGCATTCCT TTATACTGAC TTCAAGTCTT 10260
AAGACAGAGC TGAACTCTTT CAACCAGCTG CTAACTAAAG AATACCTAAA ACCCACCTGT 10320
GACTTGTAAG TCTCTGCTTT GCCATGTCCT GCCTTTTCAG GCTGACCCAA TGTATACCTT 10380
CCGTGTATTG ATTTATGATT TTTACCTACA ATTCCTGTCT TCCTGAAACA TATAAAACCA 10440
AATCATAACC CAACCACCTC AGGCACACTT TCTCAGGACC TCTTGAGACT ATTCTCCCGG 10500
CCATGGTCAT TCATATCGGC ACAGAATGAA ACCTCTTTAA AATATTTTGC AGTTTTTTTC 10560
TTTCTGTTAA CATTCCTTTC CCTTGTATCA TTGCTGTTAT TAATTTCAAG TATATATAAG 10620
CATACCTAAT TAAATACATT GTTGCTATTA TTCATTTTTG AACAAACTAT TATCTGTTAA 10680
ATCAACTAAG AATAAGACAA ATATGTTGGG TGCAGTGGTG CATGCCTATA GTCTCAGCTA 10740
CTCAGAGGCT GAGGCAGGAG GATTGCTTGA GCTCAGGAGT TTAAGACCAG CCTAGGCAAC 10800
TTAGCAAGAT CATGTCTCTT AAAAAAAAAA AAAGAAAGAA AGAAAAACAA AGTTTTAGGA 10860
GGCTGAGGCA GGAGTATCAC TTGAACCCAG GACGCAGAGG TTGCAGTGAG CCGAGATCGT 10920
GCCATTATAC TCCAGCCTGG GCAACAGAGT GAGACTCTGT CTCAAAAAAA AAAAAGAAAA 10980
GAAAAGAAAA GAAAAAAAAA GTTTTTATTT TACCTTCACT TATTCCTTCT TGGATGTTCT 11040
TCCTTTATGT AGGTACAAGG TTCTGACCTA TGTTATTTTC TTTTTCTCTA AAGAACTTCA 11100
AAAGTTTCCT GCAAGGCAGG TCTACTGGCA ATGAATTCCC TCAATTTTTG CTTGACAAAG 11160
TCTTTATTTC TGCTTCACTA TTGATGGATA ATTTCACAAG AGTGTTCCTT TTGTAGATTC 11220
ACTCTTCTTA TCCTTCCCTT CAGAAATATT CTTTGACCAA CTATTGGGTC CCAGGTACTG 11280
CACTAGAGCT TTACTTCTAG TTAATTCCCA CAGCAATTCT GAGAGGTAGG TAGGTATTAT 11340
ATTCCTAGAT GCAAACTCAG AATTCAGAAG GTAAAGTGAT GAGACTGAAG GCACACAGCA 11400
AGTAAGTGGC AGAACCTAGA TTAAAACTCA TTCTTAAAAC TTTGGCTTCC TTCTCTTTTC 11460
TTTAATGGAT TCAGTTACTT CTTCTCACCC ACTCACCTTT ATCAATTTAC ATTTCAGATA 11520
AAAGTCTTCA TGTTGGCANC TCAATGTGCC TTAACAAGAA GATGCCCACA AGAAGGGTTT 11580
GATCATCGGG ACAGCAAAGT GTCTCTTCAA GAGAAAAACT GTGAACCTGT GGTAGGTTAA 11640
GATCCTTCAT AAGGGTATTT TCATGAATGG CTGTTTTTAA CTCAAGTGAA TACAATTATT 11700
TCCATTTAAA AAGCAAGGAC AATGTGAATG TACTCATTGC CACTGAACTA TATACACCTA 11760
AAAATGGTTA AAATGGCAAC TTTTATGTGT ATTTTATGAG AATAAAAAAT AAATAATAAT 11820
AAAAAACAAG GGAAGTACAG ATATTTTCTT AATTGTGTTG TCACATACCC AGTGTTTCCA 11880
GGGTCAATAA TGAGAGCCCT ACATGTAAGA TTCAAAGGAA GAATTTAGTC CTGGATACAA 11940
TATTCTTTTA TGTTTTTAGT TATATTTGCC TTTTTAATGG ATGCAGATAT ATACAGAGGG 12000
AAGGGATAAA GTACCTATTA TTTATTGTAT AGAGCTGTGC TGTCTGATGG CTTAGCCACT 12060
AGTCACATGG TGCTATTGAA CACTTAAAAC ACAGGAGTTT GAAATAAGCA TGTATTATAA 12120
TACATATCAT ATTTCAAAAA TATTAGTATG TAGAAAAGAA GATAAATGGT TCATTAATGA 12180
TTTTTATATT GATTCACCTT GAAATAAATA TTCTGAAAAT ATTAGGTTAA ACAAAATATT 12240

TTAAGATTAA TTTTACATGT TTCTTCTTTT AAATGTAGCT ACTAGAAATT TTAAAATTAC 12300
ATATGGCTGG GCATGGTGGC TCACACCTGT AATCCCAGCA CTTCGGGAGG CCGAGGTGGG 12360
TGGATCACCT GATCTCAGGA GCTCGAGACC AGCCTGGCAA ACATGGTGAA ATCCTATCTT 12420
TACTAAAAAT ACAAAAATTA GCCAAGCGTG GTGGTGCATG CCTGTAATCC CAGCTACTTG 12480
GGACGCTGAG GCAGGAGAAT CACTTGAACC CGGGAGGTGG AGGTTGCAGT GAGCCGAGAT 12540
AGTGCCACTG CACTCCAGCC TGGGAGACAA GAGCAAAACT CCATCTCAAA AATAAATAAA 12600
TAAAATAAAA TTACATAAGT GGCTTGTACC ATATTTCTAT TGGACAGCAC TAGTACATAT 12660
ACAACACAGC ATAATGGTTG AGAGCACTGA CTCTGGAGCC AAATTACTGT GTTTGATTCT 12720
TAGCTCCACA ACTTACTAGT TGTGTGACCA TGGGCAAGCG AGTTAACCTC TCTGTGCCCC 12780
AGTTTCCCAT TCTGTAACAT GAAAATAATA AAAACACTCC CCAGAATTGT TGTGAGCATT 12840
AAATGAAGCC CTGACACATT TGTTCTGGAT ACAATATCCT CTTGTTTTAT ATTTGGTAGT 12900
ATCAATGTGC CTTTAGACAC AATTACAACG ATCTCTGTGG TAAAGATGCA ATGTATATGG 12960
TGTCTATAAA TAGCATTCAA TGATTCGTTA GTTAGGGCTT GAGACTTTTA CTGTCATGGA 13020
AAATCTAGGT ATAGCTAAGC TTTTGAGATT TTGGGAACTC CTTAACCCTA TTTTTCTCTA 13080
CTCTTGCCCC CAACAATCAG CCTATATACT TGTGAAATTT AACAATTACT TCACTGGGCA 13140
GAAATTATAT GGGAACACTT AGAAATTTCA GTCCACAGGG AAAGTATAAA TATGTTAACT 13200
ATTTTAACTT AATCCCTTCC TAGAAACACA TACACTGTTG CCAAGCCCAT ATTCTCCCTT 13260
TCTTGTTCTC ACAGTTCCCA ATGCTCCACC TGCTTATGAG AAACTCTCTG CAGAACAGTC 13320
ACCACCACCT TATTCACCTT AAGAGCCAGC GAGACACCTG AGACATGCTG AAATTATTTC 13380
TCTCACACTT TTGCTTGAAT TTAATACAGA CATCTAATGT TCTCCTTTGG AATGGTGTAG 13440
GAAAAATGCA AGCCATCTCT AATAATAAGT CAGTGTTAAA ATTTTAGTAG GTCCGCTAGC 13500
AGTACTAATC ATGTGAGGAA ATGATGAGAA ATATTAAATT GGGAAAACTC CATCAATAAA 13560
TGTTGCAATG CATGATAAAA AAAAA 13585

(2) INFORMATION FOR SEQUENCE ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 648 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

AGCTGCCGTC CGGACTCTTT TTCCTCTACT GAGATTCATC TGTGTGAAAT ATGAGTTGGC 60
GAGGAAGATC GACCTATCGG CCTAGACCAA GACGCTACGT AGAGCCTCCT GAAATGATTG 120
GGCCTATGCG GCCCGAGCAG TTCAGTGATG AAGTGGAACC AGCAACACCT GAAGAAGGGG 180
AACCAGCAAC TCAACGTCAG GATCCTGCAG CTGCTCAGGA GGGAGAGGAT GAGGGAGCAT 240
CTGCAGGTCA AGGGCCGAAG CCTGAAGCTG ATAGCCAGGA ACAGGGTCAC CCACAGACTG 300

CTGCAGGTCA AGGGCCGAAG CCTGAAGCTG ATAGCCAGGA ACAGGGTCAC CCACAGACTG 300
GGTGTGAGTG TGAAGATGGT CCTGATGGGC AGGAGATGGA CCCGCCAAAT CCAGAGGAGG 360
TGAAAACGCC TGAAGAAGAG ATGAGGTCTC ACTATGTTGC CCAGACTGGG ATTCTCTGGC 420
TTTTAATGAA CAATTGCTTC TTAAATCTTT CCCCACGGAA ACCTTGAGTG ACTGAAATAT 480
CAAATGGCGA GAGACCGTTT AGTTCCTATC ATCTGTGGCA TGTGAAGGGC AATCACAGTG 540
TTAAAAGAAG ACATGCTGAA ATGTTGCAGG CTGCTCCTAT GTTGGAAAAT TCTTCATTGA 600
AGTTCTCCCA ATAAAGCTTT ACAGCCTTCT GCAAAGAAAA AAAAAAAA 648

(2) INFORMATION FOR SEQUENCE ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

Ala Leu Ser Arg Lys Val Ala Glu Leu
5

(2) INFORMATION FOR SEQUENCE ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

Phe Leu Trp Gly Pro Arg Ala Leu Val
5

(2) INFORMATION FOR SEQUENCE ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

Thr Leu Val Glu Val Thr Leu Gly Glu Val
5 10

(2) INFORMATION FOR SEQUENCE ID NO: 56:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

Ala Leu Ser Arg Lys Val Ala Glu Leu Val
5 10

(2) INFORMATION FOR SEQUENCE ID NO: 57:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57

Ala Leu Val Glu Thr Ser Tyr Val Lys Val
5 10

(2) INFORMATION FOR SEQUENCE ID NO: 58:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acid residues

(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

Ser Thr Leu Val Glu Val Thr Leu Gly Glu Val
1 5 10

(2) INFORMATION FOR SEQUENCE ID NO: 59:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

Leu Val Glu Val Thr Leu Gly Glu Val
1 5

(2) INFORMATION FOR SEQUENCE ID NO: 60:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

Lys Met Val Glu Leu Val His Phe Leu
1 5

(2) INFORMATION FOR SEQUENCE ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

Val Ile Phe Ser Lys Ala Ser Glu Tyr Leu
1 5 10

(2) INFORMATION FOR SEQUENCE ID NO: 62:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

Tyr Leu Gln Leu Val Phe Gly Ile Glu Val
1 5 10

(2) INFORMATION FOR SEQUENCE ID NO: 63:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

Gln Leu Val Phe Gly Ile Glu Val Val
1 5

(2) INFORMATION FOR SEQUENCE ID NO: 64:
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

Gln Leu Val Phe Gly Ile Glu Val Val Glu Val
1 5 10

(2) INFORMATION FOR SEQUENCE ID NO: 65:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

Ile Ile Val Leu Ala Ile Ile Ala Ile
1 5

(2) INFORMATION FOR SEQUENCE ID NO: 66:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

Lys Ile Trp Glu Glu Leu Ser Met Leu Glu Val
1 5 10

(2) INFORMATION FOR SEQUENCE ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

Ala Leu Ile Glu Thr Ser Tyr Val Lys Val
1 5 10

(2) INFORMATION FOR SEQUENCE ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

Leu Ile Glu Thr Ser Tyr Val Lys Val Leu
1 5 10

(2) INFORMATION FOR SEQUENCE ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acids

(D) TOPOLOGY linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: M3-195.203 OR SEQ ID NO: 69:

Ile Met Pro Lys Ala Gly Leu Leu Ile
5

(2) INFORMATION FOR SEQUENCE ID NO: 70:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: M3-220.228 OR SEQ ID NO: 70:

Lys Ile Trp Glu Glu Leu Ser Val Leu
5

(2) INFORMATION FOR SEQUENCE ID NO: 71:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

Ile Asn Phe Thr Arg Gln Arg Gln Pro Ser Glu Gly Ser Ser
5 10

(2) INFORMATION FOR SEQUENCE ID NO: 72:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

Leu Phe Arg Ala Val Ile Thr Lys Lys Val Ala Asp
5 10

(2) INFORMATION FOR SEQUENCE ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

Asp Val Lys Glu Ala Asp Pro Thr Gly His Ser Tyr
5 10

(2) INFORMATION FOR SEQ ID NO: 74

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1896 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74

GCGGCGGTGG CGGAGGCGGA CACATTGGCG TGAGACCTGG GAGTACGTTG TGCCAAATCA 60
TTGCCACTTG CCACATGAGT GTAAATGATG GCGGATGCAA GTATGTCCTC TGCCGATGGG 120
AAAAGCGATT ATGGCCTGCG AAGGTGACAG CCATTATTCT GTAACTTCAG GACTTAGAAA 180
TGACTTTCGG GTGACAAGTA AAATCTTGAT CAGGAGATAC CTAGGATTTG CTTCAGTGAA 240
ATAATTGAGC CAGAACACGG TTGGCACTGA TTCTCGTTCC CCATTTAATG GGGTTTTGGT 300
CTAGTGCTTC CAAGGTTACA CTTCCAGAAA TGTCTTTTTT TTTTCACACT AAAAAAAAAA 360
AAAAGAATCA GCTGTAAAAA GGCATGTAAG GCTGTAACTC AAGGAAAGAT CTGGCAAGCA 420
GCCCTGTGAT AGTAAATTAT GGTCGTGTTC AGGGAATGCT TTCCAGCAAT TCAGTAGACA 480
GTGCTCAGCT GCAATGCAAA AGCCCAGGTC CTTGTCTTTG TCTGCCACTG GCCTCTCATG 540
CCTCAGTTTC CCCATCTGTG AAACAATGGG GATTGGACCA AATATCTGAA ATCCCATGGT 600
TATAGGCCTT CAGGATTACC TGCTGCATTT GTGCTAAAGT TTGCCACTGT TTCTCACTGT 660
CAGCTGTTGT AATAACAAGG ATTTTCTTTT GTTTTAAATG TAGGTTTTGG CCCGAACCGC 720
GACTTCAACA AAAAATAAGA GAAGAAAGGA ATATTTTCTA GCTGTGCAAA TCCTCTCCCT 780
AGAGGAAAAG TTAATTGTTG TGTTGTTTTA ATACTGTTTT TTCCCGTGTA GATTTCTGAT 840
ACTTCAATCC CCTACTCCCC CAAAACAGTT GAAGCCCAGC CCACTCTTAA TGGGCTTATT 900

CACCATTTGT GTAATTCATT AATGCTCATA ATAACCTCAT GAGAAAGCAA CTAGTTTGAT 960
TTTATGTCAG TTTGGAAGCT GAAGATCCAA ACGAGGCATT CTGTGAGATC TATGGAGAGA 1020
TTGGTACAAA CACTGAATAC ATGTAAATTA TACTCAGGGT AGACCCTATT TGTGGTTAAA 1080
ATAGGGATAT TTCCTTTTTT TTTTTTTTTT TTTTGACTGT TTCTTAATCA GTGCCATGCC 1140
AGGAAAATAG GGATGTTTCC TTCCCAGAGA TCTGTGTGTC TTTTTTCAGA AACGTCTGTG 1200
ACAGGCCCAT CAATTTTGAA ATATTTGGTT TTTGAGCCTG TCACTCTAAA CCAGCGTTTA 1260
ACGTTCAAAA GGCAAATAAC TGATGACCAG GCGGCACATT GTTCTGCTCC GTGAGTGTCT 1320
GGCACTGGGA AAGGTGTAGA TTGTCTAGAA TGACAGCAAT TCCGACGCCC CAGTCAGTCC 1380
TGCGTGATTG TGGCGAGGGC GCGTCTGGCA CCGGGAAGGT GTAGATCATC TAGAATGACG 1440
GCGATTCCGA CGCCCCGGTC AGTCCTGCGT GATTGGCGAG GGTGCATCTG TCGTGAGAAT 1500
TCCCAGTTCT GAAGAGAGCA AGGAGACTGA TCCCGCGTAG TCCAAGGCAT TGGCTCCCCT 1560
GTTGCTCTTC CTTGTGGAGC TCCCCCTGCC CCACTCCCTC CTGCCTGCAT CTTCAGAGCT 1620
GCCTCTGAAG CTCGCTTGGT CCCTAGCTCA CACTTTCCCT GCGGCTGGGA AGGTAATTGA 1680
ATACTCGAGT TTAAAAGGAA AGCACATCCT TTTAAACCAA AACACACCTG CTGGGCTGTA 1740
AACAGCTTTT AGTGACATTA CCATCTACTC TGAAAATCTA ACAAAGGAGT GATTTGTGCA 1800
GTTGAAAGTA GGATTTGCTT CATAAAAGTC ACAATTTGAA TTCATTTTTG CTTTTAAATC 1860
CAGCCAACCT TTTCTGTCTT AAAAGGAAAA AAAAAA 1896

(2) INFORMATION FOR SEQUENCE ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acid residues

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

Glu Glu Lys Leu Ile Val Val Leu Phe
5

(2) INFORMATION FOR SEQUENCE ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1554 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(ix) FEATURE:

(A) NAME/KEY: DAGE

(D) OTHER INFORMATION: Xaa is Arg when V is C or A or Gly when V is G

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

```
VGA CTG AGA CCT AGA AAT CCA AGC GTT GGA GGT CCT GAG GCC AGC CTA     48
Xaa Leu Arg Pro Arg Asn Pro Ser Val Gly Gly Pro Glu Ala Ser Leu
1               5                  10                  15

AGT TTC CGC AAA ATG GAA CGA AGG CGT TTG CGG GGT TCC ATT CAG AGC     96
Ser Arg Phe Lys Met Glu Arg Arg Arg Leu Arg Gly Ser Ile Gln Ser
            20                  25                  30

CGA TAC ATC AGC ATG AGT GTG TGG ACA AGC CCA CGG AGA CTT GTG GAG    144
Arg Tyr Ile Ser Met Ser Val Trp Thr Ser Pro Arg Arg Leu Val Glu
        35                  40                  45

CTG GCA GGG CAG AGC CTG CTG AAG GAT GAG GCC CTG GCC ATT GCC GCC    192
Leu Ala Gly Gln Ser Leu Leu Lys Asp Glu Ala Leu Ala Ile Ala Ala
    50                  55                  60

CTG GAG TTG CTG CCC AGG GAG CTC TTC CCG CCA CTC TTC ATG GCA GCC    240
Leu Glu Leu Leu Pro Arg Glu Leu Phe Pro Pro Leu Phe Met Ala Ala
65                  70                  75                  80

TTT GAC GGG AGA CAC AGC CAG ACC CTG AAG GCA ATG GTG CAG GCC TGG    288
Phe Asp Gly Arg His Ser Gln Thr Leu Lys Ala Met Val Gln Ala Trp
                85                  90                  95

CCC TTC ACC TGC CTC CCT CTG GGA GTG CTG ATG AAG GGA CAA CAT CTT    336
Pro Phe Thr Cys Leu Pro Leu Gly Val Leu Met Lys Gly Gln His Leu
            100                 105                 110
```

CAC CTG GAG ACC TTC AAA GCT GTG CTT GAT GGA CTT GAT GTG CTC CTT 384
His Leu Glu Thr Phe Lys Ala Val Leu Asp Gly Leu Asp Val Leu Leu
115 120 125

GCC CAG GAG GTT CGC CCC AGG AGG TGG AAA CTT CAA GTG CTG GAT TTA 432
Ala Gln Glu Val Arg Pro Arg Arg Trp Lys Leu Gln Val Leu Asp Leu
130 135 140

CGG AAG AAC TCT CAT CAG GAC TTC TGG ACT GTA TGG TCT GGA AAC AGG 480
Arg Lys Asn Ser His Gln Asp Phe Trp Thr Val Trp Ser Gly Asn Arg
145 150 155 160

GCC AGT CTG TAC TCA TTT CCA GAG CCA GAA GCA GCT CAG CCC ATG ACA 528
Ala Ser Leu Tyr Ser Phe Pro Glu Pro Glu Ala Ala Gln Pro Met Thr
165 170 175

AAG AAG CGA AAA GTA GAT GGT TTG AGC ACA GAG GCA GAG CAG CCC TTC 576
Lys Lys Arg Lys Val Asp Gly Leu Ser Thr Glu Ala Glu Gln Pro Phe
180 185 190

ATT CCA GTA GAG GTG CTC GTA GAC CTG TTC CTC AAG GAA GGT GCC TGT 624
Ile Pro Val Glu Val Leu Val Asp Leu Phe Leu Lys Glu Gly Ala Cys
195 200 205

GAT GAA TTG TTC TCC TAC CTC ATT GAG AGA GTG AAG CGA AAG AAA AAT 672
Asp Glu Leu Phe Ser Tyr Leu Ile Glu Arg Val Lys Arg Lys Lys Asn
210 215 220

GTA CTA CGG CTG TGC TGT AAG AAG CTG AAG ATT TTT GCA ATG CCC ATG 720
Val Leu Arg Leu Cys Cys Lys Lys Leu Lys Ile Phe Ala Met Pro Met
225 230 235 240

CAG GAT ATC AAG ATG ATC CTG AAA ATG GTG CAG CTG GAC TCT ATT GAA 768
Gln Asp Ile Lys Met Ile Leu Lys Met Val Gln Leu Asp Ser Ile Glu
245 250 255

GAT TTG GAA GTG ACT TGT ACC TGG AAG CTA CCC ACC TTG GCG AAA TTT 816

```
Asp Leu Glu Val Thr Cys Thr Trp Lys Leu Pro Thr Leu Ala Lys Phe
            260                 265                 270

TCT CCT TAC CTG GGC CAG ATG ATT AAT CTG CGT AGA CTC CTC CTC TCC   864
Ser Pro Tyr Leu Gly Gln Met Ile Asn Leu Arg Arg Leu Leu Leu Ser
        275                 280                 285

CAC ATC CAT GCA TCT TCC TAC ATT TCC CCG GAG AAG GAA GAG AGT ATA   912
His Ile His Ala Ser Ser Tyr Ile Ser Pro Glu Lys Glu Glu Ser Ile
    290                 295                 300

TCG CCC AGT TCA CCT CTC AGT TCC TCA GTC TGC AGT GCC TGC AGG CTC   960
Ser Pro Ser Ser Pro Leu Ser Ser Ser Val Cys Ser Ala Cys Arg Leu
305                 310                 315                 320

TCT ATG TGG ACT CTT TAT TTT TCC TTA GAG GCC GCC TGG ACT CAG TTG  1008
Ser Met Trp Thr Leu Tyr Phe Ser Leu Glu Ala Ala Trp Thr Gln Leu
                325                 330                 335

CTC AGG CAC GTG ATG AAC CCC TTG GAA ACC CTC TCA ATA ACT AAC TGC  1056
Leu Arg His Val Met Asn Phe Leu Glu Thr Leu Ser Ile Thr Asn Cys
            340                 345                 350

CGG CTT TCG GAA GGG GAT GTG ATG CAT CTG TCC CAG AGT CCC AGC GTC  1104
Arg Leu Ser Glu Gly Asp Val Met His Leu Ser Gln Ser Pro Ser Val
        355                 360                 365

AGT CAG CTA AGT GTC CTG AGT CTA AGT GGG GTC ATG CTG ACC GAT GTA  1152
Ser Gln Leu Ser Val Leu Ser Leu Ser Gly Val Met Leu Thr Asp Val
    370                 375                 380

AGT CCC GAG CCC CTC CAA GCT CTG CTG GAG AGA GCC TCT GCC ACC CTC  1200
Ser Pro Glu Pro Leu Gln Ala Leu Leu Glu Arg Ala Ser Ala Thr Leu
385                 390                 395                 400

CAG GAC CTG GTC TTT GAT GAG TGT GGG ATC ACG GAT GAT CAG CTC CTT  1248
Gln Asp Leu Val Phe Asp Glu Cys Gly Ile Thr Asp Asp Gln Leu Leu
```

405 410 415

GCC CTC CTG CCT TCC CTG AGC CAC TGC TCC CAG CTT ACA ACC TTA AGC 1296
Ala Leu Leu Pro Ser Leu Ser His Cys Ser Gln Leu Thr Thr Leu Ser
420 425 430

TTC TAC GGG AAT TCC ATC TCC ATA TCT GCC TTG CAG AGT CTC CTG CAG 1344
Phe Tyr Gly Asn Ser Ile Ser Ile Ser Ala Leu Gln Ser Leu Leu Gln
435 440 445

CAC CTC ATC GGG CTG AGC AAT CTG ACC CAC GTG CTG TAT CCT GTC CCC 1392
His Leu Ile Gly Leu Ser Asn Leu Thr His Val Leu Tyr Pro Val Pro
450 455 460

CTG GAG AGT TAT GAG GAC ATC CAT GGT ACC CTC CAC CTG GAG AGG CTT 1440
Leu Glu Ser Tyr Glu Asp Ile His Gly Thr Leu His Leu Glu Arg Leu
465 470 475 480

GCC TAT CTG CAT GCC AGG CTC AGG GAG TTG CTG TGT GAG TTG GGG CGG 1488
Ala Tyr Leu His Ala Arg Leu Arg Glu Leu Leu Cys Glu Leu Gly Arg
485 490 495

CCC AGC ATG GTC TGG CTT AGT GCA ACC CCT GTC CTC ACT GTG GGG ACA 1536
Pro Ser Met Val Trp Leu Ser Ala Thr Pro Val Leu Thr Val Gly Thr
500 505 510

GAA CCT TCT ATG ACC CGG 1554
Glu Pro Ser Met Thr Arg
515

We claim:

1. Composition of matter comprising:

(i) a tumor rejection antigen precursor or a tumor rejection antigen, and at least one of:

(ii) a pharmaceutically acceptable adjuvant, and a T or B cell growth factor.

2. The composition of matter of claim 1, wherein said tumor rejection antigen precursor is a MAGE protein.

3. The composition of matter of claim 1, wherein said tumor rejection antigen precursor is a BAGE protein.

4. The composition of matter of claim 1, wherein said tumor rejection antigen precursor is a GAGE protein.

5. The composition of claim 1, wherein said tumor rejection antigen is derived from a MAGE protein.

6. The composition of matter of claim 5, wherein said MAGE protein is MAGE-1, MAGE-2, or MAGE-3.

7. The composition of matter of claim 6, wherein said tumor rejection antigen consists of one of SEQ ID NO: 1 through SEQ ID NO: 5.

8. The composition of matter of claim 1, wherein said pharmaceutically acceptable adjuvant is a substantially pure saponin derived from Quillaja saponaria.

9. The composition of matter of claim 8, wherein said substantially pure saponin is selected from the group consisting of QA-7, QA-21, QA-17, and QA-18.

10. The composition of matter of claim 1, wherein said pharmaceutically acceptable adjuvant is MTP-MF59.

11. Method for stimulating an immune response in a subject comprising administering the composition of matter of claim 1 to said subject in an amount sufficient to provoke an immune response to said tumor rejection antigen precursor or tumor rejection antigen.

12. The method of claim 11, wherein said immune response comprises proliferation of T cells specific for complexes of said tumor rejection antigen and a major histocompatibility complex molecule to which said tumor rejection antigen binds.

13. The method of claim 11, wherein said T cells are cytolytic T cells.

14. The method of claim 11, wherein said immune response comprises production of antibodies against said tumor rejection antigen precursor or tumor rejection antigen.

15. The composition of claim 1, in intravenous form.

16. The composition of claim 1, in the form of a liposome.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/12463

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 38/00, 39/385, 45/00, 45/05; C07K 7/00, 14/00, 14/82
US CL :424/185.1, 193.1, 277.1; 530/300, 350, 395, 868
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/185.1, 193.1, 277.1; 530/300, 350, 395, 868

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Examiner's tumor rejection file references

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Automated Patent System and DIALOG (file = BIOCHEM). Key words: MAGE, tumor rejection antigen?

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, A, 5,342,774 (BOON ET AL) 30 August 1994.	1-16
A, P	US, A, 5,405,940 (BOON ET AL.) 11 April 1995.	1-16
A, E	US, A, 5,462,871 (BOON-FALLEUR ET AL) 31 October 1995.	1-16
Y	R. W. Ellis (ed.) "Vaccines: New Approaches To Immunological Problems", published 1992 by Butterworth-Heinemann (Stoneham, MA), pages 431-449, see entire document.	1-16
Y	Vaccine, Volume 11, Issue 3, issued 1993, Gupta et al., "Adjuvants - a balance between toxicity and adjuvanticity", pages 293-306, see entire document.	1-16

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
03 JANUARY 1996	09 FEB 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer THOMAS CUNNINGHAM
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/12463

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Immunology Today, Volume 14, No. 6, issued 1993, Hadden, "Immunostimulants", pages 275-280, see entire document.	1-16